

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 (Search time 6250.78 Seconds
(without alignments) 3295.534 Million cell updates/sec)

Title: US-09-806-955-3

Sequence: 1 atggagagagacacagaaga.....tccacacacacacacacacac

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372869281 residues

Minimum DB seq length: 0 22703874

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666.8	86.9	2539	12 AK004578	AK004578 Mus muscu
2	860.2	44.9	861	10 AL552861	AL552861 Mus muscu
3	812	42.4	825	11 BG675308	BG675308 602621612
4	812	42.4	1100	10 AL541667	AL541667 AL541667
5	796.8	41.6	814	11 BG681177	BG681177 602629315
6	770.6	40.2	777	11 BG743404	BG743404 602639114
7	754	39.3	782	11 BT260800	BT260800 602970779
8	747.4	39.0	946	10 BE740540	BE740540 601595632
9	735	38.3	736	11 BG674623	BG674623 602620671
10	731	38.1	769	10 BE742027	BE742027 601594538
11	726.6	37.9	1055	11 BE876797	BE876797 601488373
12	724.6	37.8	768	10 AI909687	AI909687 PM-BT219-

13	721.6	37.6	985	11 B1086390	B1086390 602849835
14	719.8	37.5	835	11 BG698078	BG698078 602659944
15	718.6	37.5	813	11 BG679349	BG679349 602628541
16	716.8	37.4	1205	11 BF792328	BF792328 602252860
17	711	37.1	967	10 AL555722	AL555722 AL555722
18	706.6	36.9	1003	10 AL569021	AL569021 AL569021
19	703.6	36.7	961	11 BG110618	BG110618 602281078
20	702	36.6	976	10 AL546661	AL546661 AL546661
21	681.4	35.5	874	10 BE740407	BE740407 601595436
22	672.8	35.1	775	11 BG118890	BG118890 602347893
23	668.4	34.9	778	11 BF341614	BF341614 602016039
24	667.6	34.8	838	10 BE746886	BE746886 601579391
25	665.6	34.7	918	10 AL558058	AL558058 AL558058
26	662.4	34.6	935	11 BG116226	BG116226 602318446
27	661	34.5	827	11 BG253599	BG253599 602364334
28	660.2	34.4	762	11 BG874601	BG874601 RC3-BN005
29	659.8	34.4	706	11 BP984440	BP984440 602307982
30	659.4	34.4	706	10 AU136376	AU136376 AU136376
31	659.2	34.4	829	11 BG684785	BG684785 602799210
32	659.2	34.4	919	11 BF311360	BF311360 601896761
33	656.2	34.2	923	11 BG685362	BG685362 602783969
34	655.2	34.2	807	11 B1101144	B1101144 602886829
35	654.8	34.2	829	11 BG923286	BG923286 602825247
36	649.8	33.9	766	11 BG681518	BG681518 602795482
37	649.8	33.9	923	11 BG255899	BG255899 602367542
38	648.6	33.8	857	11 B1218638	B1218638 602937865
39	647.6	33.8	924	10 AL542474	AL542474 AL542474
40	647	33.8	953	11 BG289393	BG289393 602381421
41	646.2	33.7	1019	10 AL578162	AL578162 AL578162
42	642.6	33.5	682	11 BG115553	BG115553 602317239
43	642	33.5	673	11 BG283834	BG283834 602407558
44	640.4	33.4	763	10 AV755667	AV755667 AV755667
45	639.6	33.4	893	10 BE790126	BE790126 601482941

ALIGNMENTS

RESULT 1	AK004578	2539 bp	mRNA	HTC	05-JUL-2001
LOCUS	AK004578				
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200003105, full insert sequence.				
ACCESSION	AK004578	1	GI:12835844		
VERSION	AK004578.1				
KEYWORDS	CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:1200003105.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1 (bases 1 to 2539)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Methods in enzymology. 303, 19-44 (1999)				
PubMed	98279253				
REFERENCE	10349636				
AUTHORS	2 (bases 1 to 2539)				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome research. 10 (10), 1617-1630 (2000)				
PubMed	20499374				
REFERENCE	11042159				
AUTHORS	3 (bases 1 to 2539)				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, D., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Saitoh, H., Sakaguchi, S., Ikegami, T., Kashii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,				

Query Match:	40.9%	Score 783.8:	DB 1	Length 2574:
Best Local Similarity:	65.9%	Pred. No. 1.1e-215:		
Matches 1247:	Conservative 0:	Mismatches 627:	Indels 17:	Gaps 7
QY	4	gaagaggaacaaagagagcgtlbgacagtgatgcgcatcgacctlbgagaccacac	63	
Db	519	GATCATATACCTCTACAGATTCATATGCGMACAGTTATGTTGATATGATCTTCTGTCACAACTAC	578	
QY	64	tccgcgcgcgcgcgtlgtccaaagaaagccgcgcgcgtgagagatcatcgccaaagatcaagagaaac	123	
Db	579	TCCTTGCTTGCCTTATGAAATAATGCTCTGTACAAATTTATTTGCCAACGATCAGCGTAT	638	
QY	124	cgcacagccgcctccatctatgctgcctcactcctctgaagaggaaagcgtlgtgctgac	183	
Db	639	CGTATTTACACCTCTCATATGTGCGCTT--TACTGAAAGACGAACGTTTGCTGTGTAGAGGCC	695	
QY	184	gcacaaagacagcgtccacccccaaccccgcgaagacacgcgtctltgaagcgaacgcgcctac	243	
Db	696	GCTAAGAACCAAGCTCTCTTCATCTGTAACCTGTAACACCATTTTGGATATACACGCTTAT	755	
QY	244	ggccgcgcgtggaatgaacccgcctctgtlgtcagcagagacacaaatctctgcgtlcaag	303	
Db	756	GGAGGTGAGTTTACGGAAGAACAAATGGCCAAAGGATATTTAAATCTTTTCCCTTCCATAT	815	
QY	304	gtcaaaagaagaacaaacatacatctcagatctgatltgaaggtgaggtgcgaacaagaca	363	
Db	816	GTAATATGACAAAGACCGCTCTTGGTTGAGCTTAAATGATGAGTGCTAAGACAAAGAAC--	872	
QY	364	tttgctctgaagaabaattctgcagatggtctccacaaatgaagaagaaacgcgtgagc	423	
Db	873	TTTACCCCTGACAAATTTTCAGCATGATTTCTTAGAAATGACAAACATGCTGAAGCT	932	
QY	424	tattcggaaagaaggtltaaccaatgaagtlgtltaactgtltaacagcctatltaaatgatgc	483	
Db	933	TACCTCGAAGACCTGTCACCTACTCTGTTGTTACGTCGCCGCTACTTCAATGAGCGCT	992	
QY	484	caagcgaacgaacccaaagacgcgtgaagatctatgctcgcgcctaaatgltatgaagatc	543	
Db	993	CAGCGTAGCGTACCAAGATGCTGTGTACTATTTGCGGCTTGAATGTTATTCGATGCTC	1052	
QY	544	aacgagcctacgcgcagctgtatctgtltaagcctlgtatgaagaggagggagaaagaaac	603	
Db	1053	AATGAGCGTACTGCGGCTGTATGGCTTATGGCTACGGAATTAGACAAACATGATACGAGACAT	1112	
QY	604	atccctgltttgaacctgggtgagcgaacctcgaatlgltcctctccatctgaacat	663	
Db	1113	ATTGTTGTTATGATTTAGTGTGTGTGATCTTTGACGTTTCTTTCTTATTGACAT	1172	
QY	664	gggtlctcgaagtlgtgscacaaatgaagaaatgaagaaagcggcaaaatgltcagg	723	
Db	1173	GGTGTTTTCGAAGTTTGGCTACTTACAGGATACCATCTCGGTGTGAGAGACTTGAC	1232	
QY	724	caggtgtcaltgaagaaacatcatcaaacctgttcaaaaagaagaaagcggcaaaatgltcagg	783	
Db	1233	AACCGTGTATCACTACTACTTACGCCCTACTTTCACCCGACAAACAAATGTGATGTATCT	1292	
QY	784	aaagacaatagagctgtgcagaactccgcgcgcgagltgaagaagggccaaacgcgcctg	843	
Db	1293	AAGCATTTAAGCGTATATGGGAAACCAAGCGTGAAAGTTGAAAAAAGCCAAACGCTACTTTG	1352	
QY	844	tctctcgcgcacaaagaaagaaatlgaaatltgagctctctatgaagaggaagacatcttc	903	
Db	1353	TCCTCCCAAAAGTCTGTTGCTATTCGAGATTTGAATCTTTTAAAGGTCAAAGACTTTCT	1412	
QY	904	gagacccctgaactcggcgcaaaatttgaagag-ctcaaacatggaatctgtcctgcctactac	962	
Db	1413	GAAACTTATTCCTCTGCTAAGTTTCGAGGAGGATTTAAACATGATGATCTCTTCAAGAGACTTT	1472	
QY	963	gaagacccctcagaagaaatgtctggaagaaatctgatltgaagaagctctgatctgatgaat	1022	
Db	1473	G-ACCGCTGTTCAGCAATATTAAAGACATCCAAACCTCAAGAAATCCGAGATTTGATGAT	1511	
QY	1023	tgctctctgtgtgtgcgcgcgcgaatltccaaagatctcgaacatgtgtltaagaagctct	1082	

```

Db      1552  CGTCTTGTGCGTGGTCTCTACTGCTATCCTTAAGTTCAAGAAC---TTTTCGAGACGTT 1588
Oy      1083  caatgtcgaaggaaacacatcccgctggcatalaacaccagaatgaagctctgtagctatgctgtgc 1142
Db      1589  CTTTGGTAAAGAAAGGCTTCTTAAGGCTATCAATCCCGATGAGGCGTTGTCCTATGAGGCTGC 1648
Oy      1143  tctcgaagctgtgtgtctctctctgtgtgatacagaatacaggtagaagctgtgactgtgtgtgt 1202
Db      1649  TGTTCACACCGCGCTTTTATCTGTGCGCAGAGAGCAAGCAATGTATTAATCTGCTCTTGAGACT 1708
Oy      1203  atgtccctctacacttggtatctgaacctgttggaagggtgtcaatgacccaactgtatcccaag 1262
Db      1709  TATCCCTCTTACCTTTAGGTATTTAGATACACCGGTGTGTATATGACTAAACCTTATGTCGCG 1768
Oy      1263  gaaccacagtggtgtctaccgaagaagctctgacatctctcttaagaagctctcgtataatcaacc 1322
Db      1769  TAAACACTCTATCTCTACTGCTAGTAAAGTCGCAAAATTTCTCTACTGCGGTTGACAAATCAAA 1828
Oy      1323  aactgttaacatacgaagctctatgaagggtgtgaagaaacccctgaaacaaagaaacatacttct 1382
Db      1829  TACGTGTTTAAATCAACTCTATGATGAAGGTGAAGGTCAAGTACTCTTACTAAGCAACAAACCTTCT 1888
Oy      1383  gggtacatcttgatctgtactgtgaattctctctctgtctctgtgtgtgtgtgtgtgtgtgtgt 1442
Db      1889  TGGAAATATTTCACTCTGTGTGTATTCCTCTCTGCGCCCTGTGGTGTCCCAATATGAGAT 1948
Oy      1443  cacccttgagatgaatgttgaaagtgtatctctctgtgaagtagaacgtctgaagaagaag---gtctc 1499
Db      1949  CACGTTTGAAGTGCATGCCAATAGGTGTTTGTGACTGTTTTCAGCCGTCGACAAATCTGTGTA 2008
Oy      1500  agggaaacanaaaataagaatctcaacatcaacaaatgaaccgaagatgctgtacacctgaagaat 1558
Db      2009  GCGTAAAGCTCAGAAAGCTTGTATTCAGAAATGCAAAAGGCGTTTGTCTGAGAGCAATAT 2068
Oy      1560  cgaagaagtggtatgaatgaatctgtgaagatgtgtctgtgaaggaagacaaaagctcaagaagcg 1618
Db      2069  CGAGCGCATGTTTAAAGAGGCGCAAGAAATTCCTCTGAAGAAAGATTAAGATTTTGAAGAGCG 2128
Oy      1620  catgtatcatgaaaatgaatgtgaaagctatgtcctatctctctaaagaatcagaattgaga 1679
Db      2129  TATTGAACCTGTAATACCTCTTGAACACTACGCGCTATGCTTTGCAAAAGCTCAATTTGACGA 2188
Oy      1680  taagaagaagctgtggaaggttaactctctctctgaagaataaggaagcaatgaaaaagctgt. 1739
Db      2189  TGATGAGCAATTAAGTGTGTAAGTTGATCCCGAAGATMAACAGACGTTTTCGACGCTGT 2248
Oy      1740  aagaagaagaatgtgaatgtgtgtgaagaagccac---caagaatgtcatatgaagactctaa 1796
Db      2249  CGAAGATGTGCTCAATTCGCTTGAAATCCACGAGAGAAATGCCAGCAACGAAGAAATTTGCA 2308
Oy      1797  agctgaagaagaaggaactgtgaagaagaatgtgtcaaccaatatacagaacactctatgtgaag 1856
Db      2309  AGATCAGCGTCAAAAACACTGCATGCCGTTGTTCTATTTACCCAAAATGTGTATTTCGA 2366
Oy      1857  tgcaggccctccccaactgtgtgaagaagat 1887
Db      2369  AGGACCTGCTGATGCTGATGAGAGATGAT 2399

RESULT 2
US-08-441-139-1
: Sequence 1, Application US/08441139
: Patent No. 5773245
: GENERAL INFORMATION:
: APPLICANT: Wiltrop, Dr. Karl D.
: APPLICANT: Robinson, Anne S.
: TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
: STREET: 400 Garden City Plaza

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 : Search time 617.47 seconds
(without alignments)
2661.656 Million cell updates/sec

Title: US-09-806-955-3
Perfect score: 1917
Sequence: 1 atgagagagacagaaga.....tccaccaccaccaccaccac 1917

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 segs, 428662619 residues

1 number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N.Geneseq.1101: *
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1917	100.0	1917	21	CDNA encoding huma
2	1892.4	98.7	2520	21	Human ovarian huma
3	858.4	44.8	2446	22	Human cervical can
4	810	42.3	1977	21	Arabidopsis thalia
5	805	42.0	2295	21	Arabidopsis thalia
6	757.8	39.5	1968	21	Arabidopsis thalia
7	754	39.3	2549	22	S cerevisiae apopt
8	733.8	38.4	2564	22	C albicans apopt
9	723.8	37.8	2674	21	GFP-HSC70 fusion g
10	720.6	37.6	2052	20	Mouse heat shock c
11	712	37.1	2361	21	Lung cancer associ

12	701.8	36.6	1941	20	AA211499	T. gondii antigen
13	701.8	36.6	1941	22	AA211499	Toxoplasma gondii
14	697.2	36.4	2000	18	AA258370	Candida albicans h
15	692	36.1	1941	21	AA258370	Arabidopsis thalia
16	675.8	35.3	2700	21	AA258370	Human Hsp72 (heat
17	666.8	34.8	2465	21	AA258370	Human heat shock p
18	663.6	34.6	2465	18	AA258370	Human heat shock p
19	636.6	33.2	3607	18	AA258370	Cryptosporidium pa
20	633.6	33.1	2376	11	AA003853	Hsp70 antigen gene
21	628.4	32.8	1929	22	AA258370	Murine hsp70 cDNA
22	623.8	32.5	2419	9	AA258370	Sequence of Trypan
23	619.8	32.3	2386	18	AA258370	Marmoset intracell
24	587.2	30.6	2302	21	AA258370	Human heat shock p
25	584	30.5	2302	21	AA258370	Human heat shock p
26	568	29.6	618	20	AA258370	Apoptin-associatin
27	556.2	29.0	619	20	AA258370	Apoptin-associatin
28	555.8	29.0	622	20	AA258370	Apoptin-associatin
29	551.4	28.8	2351	21	AA258370	Triphoderna reesel
30	548.4	28.6	3470	14	AA258370	Human glucose regu
31	541.4	28.2	1611	13	AA258370	Antigen mc-5c gene
32	469.4	24.5	2155	18	AA258370	Murine mortalin CD
33	446.6	23.3	2298	21	AA258370	Fusarium venenatum
34	445.2	23.2	661	20	AA258370	Apoptin-associatin
35	431	22.5	910715	20	AA258370	Borrelia burgdorfe
36	429	22.4	2515	20	AA258370	Ehrlichia sp. B3 g
37	425.8	22.2	2438	18	AA258370	S. agalactiae heat
38	425.6	22.2	1935	19	AA258370	Nitrosomonas denit
39	422	22.0	269223	22	AA258370	Genomic fragment #
40	419.6	21.9	1266	21	AA258370	Arabidopsis thalia
41	419	21.9	1620	22	AA258370	S. epidermidis ope
42	419	21.9	3488	22	AA258370	S. epidermidis gen
43	418.2	21.8	2357	21	AA258370	Arabidopsis thalia
44	417.4	21.8	1929	21	AA258370	Neisseria meningit
45	417.2	21.8	2465	21	AA258370	Neisseria meningit

ALIGNMENTS

RESULT	ID	CDNA	Location/Qualifiers
1	AAA30792	standard; cDNA; 1917 BP.	
2	AAA30792	(first entry)	
3	AAA30792	CDNA encoding human immunoglobulin heavy binding protein, BIP(78KD).	
4	AAA30792	Immunoglobulin heavy chain binding protein; BIP(78KD); chondrocyte;	
5	AAA30792	autoantigen; rheumatoid arthritis; antiarthritic; antineumatic;	
6	AAA30792	p78; ss.	
7	AAA30792	Homo sapiens.	
8	AAA30792	Key	location/Qualifiers
9	AAA30792	CD5	1..1917
10	AAA30792	FT	/*tag- a
11	AAA30792	FT	/product- "Human BIP(78KD)
12	AAA30792	FT	/note- "No stop codon given in the specification"
13	AAA30792	FT	
14	AAA30792	FT	
15	AAA30792	FT	
16	AAA30792	FT	
17	AAA30792	FT	
18	AAA30792	FT	
19	AAA30792	FT	
20	AAA30792	FT	
21	AAA30792	FT	
22	AAA30792	FT	
23	AAA30792	FT	
24	AAA30792	FT	
25	AAA30792	FT	
26	AAA30792	FT	
27	AAA30792	FT	
28	AAA30792	FT	
29	AAA30792	FT	
30	AAA30792	FT	
31	AAA30792	FT	
32	AAA30792	FT	
33	AAA30792	FT	
34	AAA30792	FT	
35	AAA30792	FT	
36	AAA30792	FT	
37	AAA30792	FT	
38	AAA30792	FT	
39	AAA30792	FT	
40	AAA30792	FT	
41	AAA30792	FT	
42	AAA30792	FT	
43	AAA30792	FT	
44	AAA30792	FT	
45	AAA30792	FT	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2001, 10:55:47 ; Search time 11004.8 Seconds
(without alignments)
2873.749 Million cell updates/sec

Title: US-09-806-955-3

Perfect score: 1917
1 atggagagagagagaaga.....tccaccaccaccaccaccac 1917

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

1 number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_in:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rtd:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rtd:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1917	100.0	1917	9	AF188611	AF188611 Homo sapi
2	1894	98.8	1965	9	AF216292	AF216292 Homo sapi
3	1894	98.8	2007	9	HS2471729	AJ271729 Homo sapi
4	1894	98.8	2639	6	AX014912	AX014912 Sequence
5	1866.8	97.4	2554	4	SSRNABIP	H87949 H. sapiens m
6	1697.8	88.6	2453	4	SSRBP78	X92446 S. scrofa mr
7	1670	87.1	2282	2	KMBIPCHAP	AJ002387 Mus muscu
8	1665.8	86.9	181429	2	ALJ54702	ALJ54702 Homo sapi
9	1662	86.7	2334	10	CRGPR	M17109 Chinese ham
10	1662	86.7	2408	2	MUSGRP784	D78645 Mouse mRNA
11	1636.4	85.4	159816	2	AC011026	AC011026 Homo sapi
12	1631.6	85.1	2383	10	RATBIP	M14050 Rat immunog
13	1457	76.0	1667	10	BC005785	BC005785 Mus muscu
14	1426.4	74.4	2389	5	CHNGR78	M27260 Chicken 78-
15	1261	65.8	2470	5	XL055059	M56807 Xenopus lae
16	1232.2	64.3	2379	5	AB016836	AB016836 Bombyx mo
17	1129.2	58.9	2643	3	AB016836	AB016836 Bombyx mo
18	1119.6	58.4	138532	2	ALJ39222	ALJ39222 Homo sapi
19	1077.6	55.2	2483	3	ACBIP	Z13041 A. callifor
20	1068.8	53.8	7298	2	AC014024	AC014024 Drosophila
21	1031	52.4	2117	3	SDGR78	Y09500 S. domincula
22	1003.6	51.5	2152	3	AY046874	AY046874 Trichinel
23	988.2	48.9	1956	3	ECCGRPA	M63605 Echinosuccu
24	938	48.6	2136	3	ECCGRPA	M63604 Echinosuccu
25	931.4	48.3	2118	3	AF044412	AF044412 Schistos
26	925.8	48.3	2626	8	PMSLBIPRT	Z49764 Pseudotsuga
27	851	44.4	2626	8	NTBLP4	X60057 Nicotiana t
28	843.6	44.0	2345	8	PCGRP78G	AJ295617 Corylus a
29	836.8	43.7	3859	8	CAV295617	X75673 P. cinnamomi
30	835.2	43.6	2007	8	NTBLP5	AJ295617 Corylus a
31	834.8	43.5	2273	8	AF031241	M60058 Nicotiana t
32	834.4	43.5	2180	8	AF031241	AF031241 Glycine m
33	834.4	43.5	2365	8	ZM058209	U58209 Zea mays lu
34	827.2	43.2	2322	8	CSA249329	AJ249329 Cucumis s
35	827	43.1	2040	8	SPHSC70A	AJ312020 Scherffel
36	826.2	43.1	2410	8	TOMBIPEGRC	L23551 Spinaclia ol
37	824.8	43.0	2232	8	ZM058208	U08830 Tomato BIP
38	823.2	42.9	2338	8	AF006825	U088208 Zea mays lu
39	821.6	42.9	2417	8	YLU063136	AF006825 Oryza sat
40	816.8	42.6	2569	8	SPAC22A12	U63136 Yarrowia li
41	810	42.3	32814	3	AF110397	Z298295 S. pombe chr
42	805.6	42.0	2738	8	AF110397	AF110397 Toxoplas
43	803.8	41.9	2028	8	BLVHSPRAA	L32165 Hordeum vul
44	803.4	41.9	2220	8	ATFBIPA	D84414 Arabidopsis
45	799	41.7	2194	8	PTU29675	U29675 Phaeodactyl

ALIGNMENTS

RESULT 1
AF188611
LOCUS Homo sapiens BIP protein (HSPAC) mRNA, partial cds.
DEFINITION AF188611
ACCESSION AF188611
VERSION AF188611.1 GI:6470149
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1917)
Fife,M.S.
BIP: an Autoantigen associated with Rheumatoid Arthritis
TITLE Unpublished
JOURNAL 2 (bases 1 to 1917)
REFERENCE Fife,M.S.
Direct Submission
AUTHORS
TITLE
JOURNAL
AUTHORS
JOURNAL

Submitted (21-SEP-1999) Molecular Immunogenetics, King's College,
Guy's Hospital, St. Thomas' Street, London SE1 9RT, UK

```

FEATURES
Source
Location/Qualifiers
1..1917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="chondrocytes"
/tissue_type="articular cartilage"
/note="RT-PCR primer pair for cloning the relevant B1p
fragment: 5' TATACATGTCAGGAGGACAGAGAGGAGC 3' Forward
5' CCACCTCGAATTCCTGCTATCCTCTTACCA 3' Reverse"
1..>1917
/gene="HSPA5"
/note="GNP78"
1..>1917
/gene="HSPA5"
/function="ER chaperonin"
/note="glucose-regulated protein (grp78)"
/codon_start=1
/product="BiP protein"
/protein_id="AAFI3605.1"
/db_xref="GI:6470150"
/translation="*MEEDKEDVGVTVGLDGLTVSCVGFKNRVEIANDGNRRTT
PSYVAPEGERLIGDAKNQIINSPENTFEDAKRLIGRTWDPVSVOODIFLFFKVV
EKRTAPYLDVIGGGQTTFAPEEISAMVLTKEETAEEALYLGKRTYHNVVPAPFND
ARQAKITDAGTACIAGLNVRGIINETPAALIVYGLDRKEGKNTLVLDGSGFTVPSL
INGVEFVAATNGDTHLWGRIDFQDORVMEHRIKYKKTKGDVKNRNVOKRIREVEK
KASDIISOARIEIESIEEGEDSEETLTAKKEEALNDLFTSKAPQKVLQESDGL
KSDIDETIOLVNGSTRIRPKIQLVKEPNCKEBSRGINDEAVAGAAVAGVLSGDDQ
TDDIVLLVQPLTGLGIEFGYQWTKLIPRNVYPTKSSITSTASDNPPIYVLIYBS
EPLRLKDHNLLETDLDTGSIIPAPRGVQIETVTEIDVNGILNVTAEKGTKKKRTIT
TNDQRLNLTPEETERNVDAEKFAEDKKLERLDTNNELESAYSLKNQIDKEKELGG
KLSSEDDKTEMAVEAEKIEWLESQDADIEDFAKKKLEELIYVPIKSLVSGAGPP
TGEEDTAELHHHHHH"
577 a 419 c 487 g 434 t
BASE COUNT
ORIGIN

```

Query Match	100.0%	Score 1917:	DB 9:	Length 1917:
Best Local Similarity	100.0%	Fred. No. 0:		
Matches 1917:	Conservative	0:	Mismatches	0:
			Indels	0:
			Gaps	0:
OY 1	atgagagagagacagaagagagagcgtgagcagcagcgtgcgcacatcgacccgctggaccac	60		
Db 1	ATGAGAGAGAGACAAGAGAGAGCGTGGGCGACGGTGGGTGGCATCGACCTGGGGACGACC	60		
OY 61	tatccctcggtcggcgtgttccaagaacgcgcgctgtagatcatcgcgcaacgatcagggc	120		
Db 61	TATCCCTCGGCTCGGCGTGTTCGAAGACGGCGCGTGAGATCATCGCAACGATCAGAGCG	120		
OY 121	aacgcagatcacgcgcgttcctatgtcgacctcactcctgtgaagagagacgtctcgatctggcat	180		
Db 121	AACGCATCACGCGCTCATGTGCGGCTTCACTCCTGAAGAGGAACGTGTGATTGGGAT	180		
OY 181	ggcgcccaagaacagatctcaacctcaaaacccgcgagaaacagcgtcttgacgcgaagcggctc	240		
Db 181	GGCGCCCAAGAACAGGCTCACTCAACCCCGAAGACGGCTTTGTGACGCCAAGCGGCTC	240		
OY 241	atcggcgcagcagctggaatgacccgctctgtgcagcagagacacaaagtctcttgcgcttcaag	300		
Db 241	ATCGCGCGCAGCTGGAAATGACCGGCTGTGCGAGCGACATCAAGTCTTGGCGTTCAAG	300		
OY 301	gtgtgttaaaagaagaaactaaacatatacttaagtttgtatgtgaagtgaggcaaacaaag	360		
Db 301	GTGTGTAAAGAAAGAACTAAACCATTCATTCAGTTGATTTGGAGGTGGGCAAAACAAG	360		
OY 361	aacattgctcctggaagaaattctgacatggttcacataaataagaaagaaacgcgtctgag	420		
Db 361	ACATTGCTCCTCGAAGAAATTTCTGCATGTGTTCTCACTAAATGAAAGAACCGGTGAG	420		
OY 421	gctatcttgggaagaagaggtttcccatgtgaagttgttactgtacacagcctattttaatgct	480		
Db 421	GCTATTTTGGGAAGAGGTTACCCATGTGAGTTGTTACTGTGACAGGCTATTTTAAATGAT	480		
OY 481	ggcccaagcccaagcaaacccaagcagctctggaataatctgcctggccctaaatgltatagagatc	540		

Dd	481	GCCCAAGCCCAAGAACACAAAGACGCTGGAAACTATTGTGGCGCTAAATGTTATGAAGTTC	540
Oy	541	a tcaaa g a g c c t t a e g a g c a g c t g t a t c t g t a t c g c t g t a t a a g a g g a g g y a a g a a g	600
Dd	541	A T C A A G A G C G C T A C G C A G C A G C T G T A T T G C T A T T A G C G C T G A T A A A G A G G A G G A G A A G	600
Oy	601	a a a l u c c u g u g l t t g a c c c g g s l b g c g g a a c c t o g n a t g f r l c t c t c t a c a t t y a c	660
Dd	601	A M A C T C C G G G T T T G A C C T C G S G T G C G G A C C T T G A T G T C T C T T C T A C C A T T G A C	660
Oy	661	a a t g t l c t c t g a a g t l t g y c c c a t a a t g a t a c t a c t l c t g g t g y g a a g a c t t c	720
Dd	661	A A T G T G C T T C G A A G T T T G G C C A C T A A T G A G N A C T A C A C T G C G T G A A G A A G A C T T T	720
Oy	721	g a c c a g g t i f r a t g y g a a c c t t c a c a a c g t a c a a a a a a a a g g y a a a g a t g t c	780
Dd	721	G A C C A G G T G C A T G A G A A C C T T C A T C A A C T G T A C A A A A A A A A C G G A A A G A T G T C	780
Oy	781	a g g a a a g a c a a t a g a g t g t g c a g a a a c t c c g y c g y a g t a g a a a g g c c a a c c y g g c	840
Dd	781	A G G A A A G A C A T A A G C T G T G C A A A A C T C G C G C G A G A G T G A A A A G C C A A A C G G C C	840
Oy	841	c t g t c t t c t a g a c a t a a g a a a l t g a a a t l g a d t l g a t c c t a t a g a a g g a a g a c t t	900
Dd	841	C T G T C T T C T A G A C A T C A A G A A A T T G A A T T G A T G A T C C T T A T G A A G A G A A G A C T T	900
Oy	901	t c t g a g a c c t g a c t c g g g c a a a l t t g a a g a g t c a a a t g a t c t g t l c c g t c t a c t	960
Dd	901	T C T G A G A C C C G A C T C G G G C C A A T T T G A A A G A C C T C A A C T G A T C T C T C G G T C T A C T	960
Oy	961	a t g a a g c c c g t c c g a a a g g t l t g a a a g a t t c g a t t t b a a a a g t c c g a t a t g a t g a a	1020
Dd	961	A T G A A G C C C G T C C A G A A A G T T T G A A A T T C G A T T T G A A A C T G A T T A T T G A T G A A	1020
Oy	1021	a t c g t t c t t g t t g t g t c t c g a c t c g a a t t c c a a g a t t c a g c a a c t g t t a a a g a t t c	1080
Dd	1021	A T T G T T C T T G T G G T G G C T G A C T G C A A T T T C A A A A T T C A C A A C T G T T A A G A G T C	1080
Oy	1081	t t c a a t g c a a g y a a g a c c a t c c g t g y c a t a a a c c c a g e t g a a g c t g a g c t a t g t g t c	1140
Dd	1081	T T C A A T G C A A G A A A C C A T C C C G G C A T A A A C C A G A T G A A G C T G T A G C G A T T A G T G C T	1140
Oy	1141	g c t g t c a g a g c t g t g t g t c t c t g t g t a t c a a g a t a a a a g y g a c c t g a t c t g t a t	1200
Dd	1141	G C T G T C A G G C T G G T G T G C T C T G G T G A T C A A G A T A C A G G A C C T G T A C T G C T T G A T	1200
Oy	1201	g t a t c t c c c c t t a a c t t g a t t g a a a c t g t g y g a g t g t a g a t a a c a a c t g a t t c a	1260
Dd	1201	G T A T C C C C C T T A C A C T T G T A T T G A A A C T G T G G A G G T G C A T A C A A A C T G A T T C C A	1260
Oy	1261	a g g a a c a a g y g u g c t t a c a a a a g t c c a g a t c t t t c a a a g c t t c g a t a a t a a	1320
Dd	1261	A G G A A C A A G U G G U G C C T A C A A A A A T C T C A A T T T T T A C A G C T T C G A T A T A T C A A	1320
Oy	1321	c c a a c t g t l a a a t c a a g t c t a t g a a g t g a a a g a c c c t g a c a a a a g a a t a t c t t	1380
Dd	1321	C C A A C T G T T A C A A T C A A G A G C T A T A A A G G A G A A G C C C C T A C A A A A A G A C A A T C A T T T	1380
Oy	1381	c t g g t g a c a t t g t t g c a t g a t g a a t t c c c c g t c c c c t g y g g t t c c a a a t g a t g a	1440
Dd	1381	C T G G T G A A T T T G A T T G A C T G C A A T T C C T C G C T C C T G G G G T C C A C A G A T T G A A	1440
Oy	1441	g l c a c c t t t g a a t a g a t g t g a a t g t a t t c t c t g a t g a c a g c t g a a g a c a a g t a c a	1500
Dd	1441	G T C A C C T T T G A A G A T A G A T G A A T G A A T G A T T C T T C G A G T C A C A G C T G A A G A C A A G G T G A C A	1500
Oy	1501	g g g a a a a a a a t a a g t c a a a t a c a a t a a g a a g a c a a a t c g c t g a a c c t g a a a a a t c	1560
Dd	1501	G G G A C A A A A A A T A A G A T C A A A T A C C A A T A G A C A A A T G C C T A C A C C T G A A C A A T T C	1560
Oy	1561	g a a a g a t g y g t t a a t g a t g c t g a a g a t t g t c t g a g a a g a a a a a a g c t c a a a g a g y c	1620

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:27:51 ; Search time 18.76 Seconds

(without alignments)
1237.145 Million cell updates/sec

Title: US-09-806-955-2

Sequence: 1 MEDKKEDVGVYGVIGLGT.....SKLYGSAGPPTGEEDTAL 633

Scoring table: BLOSUM62
Gapop 10.0 , gapext 0.5

Searched: 100059 seqs, 36664827 residues

Minimum number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3216	99.7	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	99.5	654	1 GR78_MESAU	P07828 mesocricetu
3	3205	99.4	655	1 GR78_MOUSE	P20029 mus musculu
4	3200	99.2	654	1 GR78_RAT	P06761 rattus norv
5	3168	98.2	652	1 GR78_CHICK	Q90593 gallus galli
6	3069	95.2	658	1 GR78_XENLA	Q91883 xenopus lae
7	2722.5	84.4	667	1 GR78_APLCA	Q16956 aplysia cal
8	2708	84.0	656	1 HS7C_DROME	P29844 drosophila
9	2594.5	80.4	661	1 HS7C_CAEEL	P27420 caenorhabdi
10	2354.5	73.0	667	1 BIP4_TOBAC	Q03684 nicotiana t
11	2346.5	72.8	668	1 BIP5_TOBAC	Q03685 nicotiana t
12	2332.5	72.3	663	1 BIP3_MAZE	Q24581 zea mays (m
13	2328.5	72.2	666	1 BIP_LARCS	P49118 lycopersico
14	2324.5	72.1	668	1 BIP_SPIOL	Q42434 spinacia ol
15	2322.5	72.0	663	1 BIP2_MAZE	P24067 zea mays (m
16	2298.5	71.3	669	1 BIP1_ARATH	Q39043 arabidopsis
17	2296	71.2	668	1 BIP2_ARATH	P78695 arabidopsis
18	2227.5	69.1	662	1 GR78_NEUCR	P36604 schizosacch
19	2216	68.7	663	1 GR78_SCHPO	Q09170 yarrowia li
20	2191	67.9	670	1 GR78_TARLI	P16474 taraxacum
21	2173	67.4	682	1 GR78_YEAST	P10591 plasmodium
22	2100	65.1	655	1 GR78_PLAFO	P10591 plasmodium
23	2085.5	64.7	641	1 HS71_YEAST	P11142 homo sapien
24	2083	64.6	646	1 HS7C_HUMAN	P08109 bos taurus
25	2083	64.5	646	1 HS7C_MOUSE	P19120 bos taurus
26	2079	64.5	650	1 HS7C_BOVIN	P53421 pichia angu
27	2078	64.4	644	1 HS71_PICAN	P36413 dictyosteli
28	2076.5	64.3	640	1 HS7C_DICDI	P41797 candida alb
29	2074.5	64.3	655	1 HS71_CANAL	P19378 ciretulus
30	2072	64.3	638	1 HS7C_CRIGR	P10592 saccharomyc
31	2064	64.2	636	1 HS72_YEAST	P22010 kluyveromyc
32	2065.5	64.0	679	1 GR78_KLULA	Q01877 puccinia gr
33	2065	64.0	648	1 HS71_PTCGR	

34	2064	64.0	641	1 HS74_YEAST	P22202 saccharomyc
35	2063.5	64.0	649	1 HS73_YEAST	P09435 saccharomyc
36	2060.5	63.9	646	1 HS70_NEUCR	O01233 neurospora
37	2060	63.9	643	1 HS71_SCHPO	O10265 schizosacch
38	2059	63.8	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	63.7	651	1 HS7C_PETRY	P09189 petunia hyb
40	2052	63.6	652	1 HS7D_MANSE	Q09639 manduca sex
41	2049	63.5	651	1 HS70_ONCMY	P08108 oncorhynch
42	2046	63.4	641	1 HS72_BOVIN	O27965 bos taurus
43	2045	63.4	651	1 HS71_ARATH	P22953 arabidopsis
44	2043.5	63.4	639	1 HS72_HUMAN	P54652 homo sapien
45	2039.5	63.2	641	1 HS71_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	654 AA.
GR78_HUMAN				
AC	P11021; Q9NPF1;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN			
DE	HEAVY CHAIN BINDING PROTEIN) (BIP) (ENDOPLASMIC RETICULUM LUMENAL CA2+			
DE	BINDING PROTEIN GRP78).			
GN	HSP45 OR GRP78.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
NCBI_Taxid	9606;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-80283347; Pubmed-2840249;			
RA	Ting J., Lee A.S.;			
RT	Human gene encoding the 78,000-dalton glucose-regulated protein and			
RT	its pseudogene: structure, conservation, and regulation.*;			
RL	DNA 7:275-286(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE-cervix carcinoma;			
RL	Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE-Fibroblast;			
RT	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;			
RT	*Grp78 is involved in the quality control of the LDL-receptor.*;			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Bernandez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;			
RT	*Sequence differences between human and human BIP gene is required for			
RT	and previously reported human sequences.*;			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 1-25 FROM N.A.			
RA	MEDLINE-91126087; Pubmed-1480470;			
RA	Chao C.C.K., Lin-Chao S.;			
RT	*A direct-repeat sequence of the human BIP gene is required for			
RT	A3187-mediated inducibility and an inducible nuclear factor			
RT	binding.*;			
RL	Nucleic Acids Res. 20:6481-6485(1992).			
RN	[6]			
RP	SEQUENCE OF 22-38.			
RA	TISSUE-breast carcinoma;			
RA	MEDLINE-97293304; Pubmed-9150946;			
RA	Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.;			
RA	Simpson R.J., Dorow D.S.;			
RT	*Two-dimensional electrophoretic analysis of human breast carcinoma			
RT	proteins: mapping of proteins that bind to the SH3 domain of mixed			
RT	lineage kinase MK2.*;			

```

RL Electrophoresis 18:588-598(1997).
RP [7]
RP SEQUENCE OF 19-40.
RC TISSUE-Colon carcinoma;
RC MEDLINE=97295306; PubMed=9150948;
RA J.I.H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RA "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M19645; AA52614.1; -
DR EMBL: X87949; CA61201.1; -
DR EMBL: AJ271729; CAB71335.1; -
DR EMBL: AF216292; AAF42836.1; -
DR EMBL: X58969; CAA2595.1; -
DR PIR: A29821; A29821.
DR HSSP: P19120; IATR.
DR SWISS-2DPAGE: P1021; HUMAN.
DR MIM: 138120; -
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70.1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00014; ER_TARGET.1.
DR PROSITE: PS00329; HSP70_1; 1.
DR PROSITE: PS01036; HSP70_3; 1.
DR ATP-binding; Endoplasmic reticulum; signal.
KW SIGNAL
FT CHAIN 1 18
FT SITE 19 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 K -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
FT SEQUENCE 654 AA; 72333 MW; 59B7DB85BC32A00 CRC64;
SQ

```

Query Match 99.7%; Score 3216; DB 1; Length 654;
 Best Local Similarity 100.0%; Pred. No. 2, 5e-152;
 Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 302 ETLTRAFEEELNDLPRSTKPVOKVLEDSDKSDIDEIVLVGGSTRIPIKTOOLVKEFF 361
DB 320 ETLTRAFEEELNDLPRSTKPVOKVLEDSDKSDIDEIVLVGGSTRIPIKTOOLVKEFF 379
OY 362 NGKPSGSGINPDEAVVGAAGVAGVSGDQDGVLLDVCPLTGLTETGVGVATKILPR 421
DB 380 NGKPSGSGINPDEAVVGAAGVAGVSGDQDGVLLDVCPLTGLTETGVGVATKILPR 439
OY 422 NTVPPTKKSQIFPSASNDQPTVYIKYEGEERLTNDNHLTGFDTLGIIPAPRVPDIEV 481
DB 440 NTVPPTKKSQIFPSASNDQPTVYIKYEGEERLTNDNHLTGFDTLGIIPAPRVPDIEV 499
OY 482 TFEIDVNGILRVTAEDKGTGNKKRITITNDQRLTPEETIERVYNDAEKFAEEDKKLERI 541
DB 500 TFEIDVNGILRVTAEDKGTGNKKRITITNDQRLTPEETIERVYNDAEKFAEEDKKLERI 559
OY 542 DTRNELESYVSLKNOIGDKKGLTSSDEKTEKKAVERIEVLESHQADIEDPKAK 601
DB 560 DTRNELESYVSLKNOIGDKKGLTSSDEKTEKKAVERIEVLESHQADIEDPKAK 619
OY 602 KKELEIIVOPITISKLYGSAGPPTGEEDTAEE 632
DB 620 KKELEIIVOPITISKLYGSAGPPTGEEDTAEE 650

```

RESULT 2
 GR78_MESAU STANDARD; PRT; 654 AA.
 AC P07823;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN
 DE HEAVY CHAIN BINDING PROTEIN) (BIP).
 GN HSP45 OR GRP78.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_Taxid=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=87305586; PubMed=3623104;
 RA Ting J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
 RT "The nucleotide sequence encoding the hamster 78-kDa
 RT glucose-regulated protein (GRP78) and its conservation between
 RT hamster and rat.";
 RL Gene 55:147-152(1987).
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
 CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17169; AA51448.1; -
 DR PIR: A27414; A27414.
 DR HSSP: P19120; IATR.
 DR InterPro: IPR000886; ER_target.
 DR InterPro: IPR001023; HSP70.
 DR Pfam: PF00012; HSP70.1.
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00014; ER_TARGET.1.
 DR PROSITE: PS00329; HSP70_1; 1.
 DR PROSITE: PS00329; HSP70_2; 1.
 DR PROSITE: PS01036; HSP70_3; 1.
 DR ATP-binding; Endoplasmic reticulum; signal.
 KW

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:23:14 ; Search time 23.81 Seconds

(Without alignments)
598.261 Million cell updates/sec

Title: US-09-806-955-2

Perfect score: 3225
1 MESDKKEDVGIYVGIDLGTT.....SKLYGSAGAPPTGEEDTAEL 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

al number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	99.5	654	1	US-08-441-139-11 Sequence 11, Appl
2	3168	99.2	666	1	US-08-441-139-16 Sequence 16, Appl
3	2173	67.4	682	1	US-08-441-139-2 Sequence 2, Appl
4	2156	66.9	663	1	US-08-441-139-7 Sequence 7, Appl
5	2077	64.4	646	1	US-08-441-139-14 Sequence 14, Appl
6	2065.5	64.0	679	1	US-08-441-139-5 Sequence 5, Appl
7	1938	60.1	643	4	US-08-797-358B-3 Sequence 3, Appl
8	1547.5	48.0	679	1	US-08-214-583-2 Sequence 2, Appl
9	1474	45.7	649	1	US-08-441-139-4 Sequence 4, Appl
10	1461.5	45.3	649	1	US-09-066-047-5 Sequence 5, Appl
11	1444	44.8	607	2	US-08-472-534-5 Sequence 5, Appl
12	1303.5	40.4	600	6	5240706-1 Patent No. 5240706
13	1140.5	35.4	339	2	US-08-928-692-52 Patent No. 5196523-13
14	941	29.2	187	6	US-08-441-139-10 Patent No. 5196523
15	84	25.6	168	1	US-08-441-139-10 Patent No. 5196523
16	818.5	25.4	315	1	US-08-257-073-7 Sequence 7, Appl
17	750.5	23.3	471	1	US-08-203-905B-2 Sequence 2, Appl
18	726.5	22.5	472	1	US-08-203-905B-14 Sequence 14, Appl
19	642.5	19.9	129	6	5196523-10 Patent No. 5196523
20	607.5	18.8	999	3	US-08-770-301A-3 Sequence 3, Appl
21	607.5	18.8	999	3	US-09-175-581-3 Sequence 3, Appl
22	598	18.5	999	2	US-08-770-301A-1 Sequence 1, Appl
23	598	18.5	999	2	US-08-175-581-1 Sequence 1, Appl
24	528.5	16.4	560	2	US-08-928-692-53 Sequence 53, Appl
25	437.5	15.6	374	2	US-08-928-692-51 Patent No. 5196523
26	402	12.5	77	6	5196523-7 Patent No. 5196523
27	395	12.2	79	6	5196523-11 Patent No. 5196523

28	367.5	11.4	599	4	US-09-080-983-9 Sequence 9, Appl
29	358	11.1	549	2	US-08-770-544-6 Sequence 6, Appl
30	345.5	10.7	80	1	US-08-464-164-4 Sequence 4, Appl
31	345.5	10.7	80	1	US-08-338-057-4 Sequence 4, Appl
32	345.5	10.7	80	2	US-08-668-416-4 Sequence 4, Appl
33	323	10.0	554	4	PCT-US94-06430-7 Sequence 7, Appl
34	320	9.9	554	4	US-08-591-468-7 Sequence 7, Appl
35	267.5	8.3	714	2	US-08-472-534-3 Sequence 3, Appl
36	264	8.2	56	6	5196523-12 Patent No. 5196523
37	230	7.1	46	6	5196523-8 Patent No. 5196523
38	211	6.5	42	6	5196523-9 Patent No. 5196523
39	192	6.0	38	6	US-08-973-662-8 Patent No. 5196523
40	168.5	5.2	1786	4	US-08-235-836C-68 Sequence 68, Appl
41	147	4.6	693	4	US-07-853-913-2 Sequence 2, Appl
42	143.5	4.4	1805	1	US-09-308-375-2 Sequence 2, Appl
43	142.5	4.4	2285	1	US-07-720-589-2 Sequence 2, Appl
44	138	4.3	700	1	US-08-785-190-2 Sequence 2, Appl
45	138	4.3	700	2	US-08-785-190-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-441-139-11
Sequence 11, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/089, 997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-11
Query Match 99.5% Score 3209, DB 1; Length 654;
Best Local Similarity 99.7% Pred. No. 4,1e-259;
Matches 629; Conservative 0; Indels 0; Gaps 0;

```

QY      2 BEDKKEDVGVVGDIDLTGTTSCVGVFNKNGVVEIANDQGNRTTPTSVAFTPGEGELIGDA 61
      20 BEDKKEDVGVVGDIDLTGTTSCVGVFNKNGVVEIANDQGNRTTPTSVAFTPGEGELIGDA 79
QY      62 AKNQLTSPNPTVDARLIGRTWNPDSVOODIKPLPFVVEKTKTPYIOVDIGGGQTKT 121
      80 AKNQLTSPNPTVDARLIGRTWNPDSVOODIKPLPFVVEKTKTPYIOVDIGGGQTKT 139
Db      122 FAPETISAMVLTAKKETAELGKKTTHAVVTPAYFNAQOARKDAGTAGLWVMRII 181
      140 FAPETISAMVLTAKKETAELGKKTTHAVVTPAYFNAQOARKDAGTAGLWVMRII 199
QY      182 NEPTAAAIAYGLDKREGEKNILVFEDLGCGTFDVSLLTINDGVFEVATNGDTHLGCEDFD 241
      200 NEPTAAAIAYGLDKREGEKNILVFEDLGCGTFDVSLLTINDGVFEVATNGDTHLGCEDFD 259
QY      242 QRVMEHFILKLTAKKTKGDKVRDNRAVQKLRREVAKKRALSSOHOARIEISFEFGEDFS 301
      260 QRVMEHFILKLTAKKTKGDKVRDNRAVQKLRREVAKKRALSSOHOARIEISFEFGEDFS 319
Db      302 ETLTBAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIPKIQOLVKEFF 361
      320 ETLTBAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIPKIQOLVKEFF 379
QY      362 NCKEPRSGINPDEAVAYGAAGVAVGVLSSGDDPTGDLVLDVCPLTIGIEVGVGMKLIPIR 421
      380 NCKEPRSGINPDEAVAYGAAGVAVGVLSSGDDPTGDLVLDVCPLTIGIEVGVGMKLIPIR 439
QY      422 NTVPVETKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFDITGIPPARGVPOIEV 481
      440 NTVPVETKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFDITGIPPARGVPOIEV 499
Db      482 TFEIDVNGILVTAEDDGTGNKNTTTTNDQNLTPETIERVWNAEKPAEEDKKLKRI 541
      500 TFEIDVNGILVTAEDDGTGNKNTTTTNDQNLTPETIERVWNAEKPAEEDKKLKRI 559
QY      542 DTRNELESYASLKNQIGDKKLGKLSSEDKETMEKAVEEKIEMWLESHQADIEDFKAK 601
      560 DTRNELESYASLKNQIGDKKLGKLSSEDKETMEKAVEEKIEMWLESHQADIEDFKAK 619
Db      602 KKELEEIYQPIISKLVSAGPPTGEEEDTAE 632
      620 KKELEEIYQPIISKLVSAGPPTGEEEDTAE 650

```

RESULT 2

US-08-441-139-16

; Sequence 16, Application US/08441139

; Patent No. 5773245

; GENERAL INFORMATION:

; APPLICANT: Mittrup, Dr. Karl D.

; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: NY

; COUNTRY: USA

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/441,139

; FILING DATE: 15-MAY-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGILLO, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-139-16

```

Query Match

98.28; Score 3168; DB 1; Length 666;

Best Local Similarity 97.88; Pred. No. 1,1e-255;

Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

```

QY      2 BEDKKEDVGVVGDIDLTGTTSCVGVFNKNGVVEIANDQGNRTTPTSVAFTPGEGELIGDA 61
      32 DEKKEDVGVVGDIDLTGTTSCVGVFNKNGVVEIANDQGNRTTPTSVAFTPGEGELIGDA 91
Db      62 AKNQLTSPNPTVDARLIGRTWNPDSVOODIKPLPFVVEKTKTPYIOVDIGGGQTKT 121
      92 AKNQLTSPNPTVDARLIGRTWNPDSVOODIKPLPFVVEKTKTPYIOVDIGGGQTKT 151
QY      122 FAPETISAMVLTAKKETAELGKKTTHAVVTPAYFNAQOARKDAGTAGLWVMRII 181
      152 FAPETISAMVLTAKKETAELGKKTTHAVVTPAYFNAQOARKDAGTAGLWVMRII 211
Db      182 NEPTAAAIAYGLDKREGEKNILVFEDLGCGTFDVSLLTINDGVFEVATNGDTHLGCEDFD 241
      212 NEPTAAAIAYGLDKREGEKNILVFEDLGCGTFDVSLLTINDGVFEVATNGDTHLGCEDFD 271
QY      242 QRVMEHFILKLTAKKTKGDKVRDNRAVQKLRREVAKKRALSSOHOARIEISFEFGEDFS 301
      272 QRVMEHFILKLTAKKTKGDKVRDNRAVQKLRREVAKKRALSSOHOARIEISFEFGEDFS 331
Db      302 ETLTBAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIPKIQOLVKEFF 361
      332 ETLTBAKFEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIVLVGSTRIPKIQOLVKEFF 391
QY      362 NCKEPRSGINPDEAVAYGAAGVAVGVLSSGDDPTGDLVLDVCPLTIGIEVGVGMKLIPIR 421
      392 NCKEPRSGINPDEAVAYGAAGVAVGVLSSGDDPTGDLVLDVCPLTIGIEVGVGMKLIPIR 451
Db      422 NTVPVETKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFDITGIPPARGVPOIEV 481
      452 NTVPVETKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFDITGIPPARGVPOIEV 511
QY      482 TFEIDVNGILVTAEDDGTGNKNTTTTNDQNLTPETIERVWNAEKPAEEDKKLKRI 541
      512 TFEIDVNGILVTAEDDGTGNKNTTTTNDQNLTPETIERVWNAEKPAEEDKKLKRI 571
Db      542 DTRNELESYASLKNQIGDKKLGKLSSEDKETMEKAVEEKIEMWLESHQADIEDFKAK 601
      572 DTRNELESYASLKNQIGDKKLGKLSSEDKETMEKAVEEKIEMWLESHQADIEDFKAK 631
QY      602 KKELEEIYQPIISKLVSAGPPTGEEEDTAE 632
      632 KKELEEIYQPIISKLVSAGPPTGEEEDTAE 662

```

RESULT 3

US-08-441-139-2

; Sequence 2, Application US/08441139

; Patent No. 5773245

XX This sequence represents a human immunoglobulin heavy chain binding
 CC protein, BIP(78KD), having a 633 amino acid sequence. The invention
 CC also encompasses a BIP(78KD) protein of 639 amino acids (Y90693).
 CC The cDNA encoding BIP(78KD), also referred to as p78 in the
 CC specification, was isolated from human chondrocytes (the specialised
 CC cells of articular cartilage) and human chondrosarcoma cell lines. The
 CC BIP(78KD) cDNA of this invention contains a number of differences
 CC compared with the published sequence (Genbank accession number X87949),
 CC and has therefore been deposited with Genbank with the accession
 CC number AF188611). These differences comprise 6 single nucleotide
 CC substitutions and a codon insertion, and result in three amino acid
 CC substitutions. The BIP(78KD) proteins react with antibodies present in the
 CC serum of rheumatoid arthritis patients, and is therefore a putative
 CC autoantigen for this autoimmune disease. BIP(78KD) is also able to
 CC selectively proliferate synovial T-cells from patients with rheumatoid
 CC arthritis. BIP(78KD) or peptides derived from the protein are useful as
 CC reagents to indicate the presence of rheumatoid arthritis, and can be
 CC used in prognostic or diagnostic tests of body fluids for rheumatoid
 CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
 CC blotting. The protein or the cDNA encoding it can also be used to test
 CC for rheumatoid arthritis by detecting antibodies to the protein.
 CC BIP(78KD), its peptides and polynucleotides are also useful
 CC therapeutically.

XX Sequence 633 AA:

Query Match 100.0%; Score 3225; DB 21; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2.3e-208;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEDKKEVGVTVVDTIDCTTSCGVFKKRVRIINADGNRTSYAFTPEGRLGD 60
 DB 1 meedkkevgtvvgldtcttscgvfkngrvrlindngnrtspyaftpegrligd 60
 QY 61 AAKNOLTSNPENTVEDAKRLIGRTWNPVODIKFLPKVVEKTKPYIQVDIGSGQTK 120
 DB 61 aaknqltsnpentvedakrligrtwnpvoaikflpkvvektkpyiqvdigsgqtk 120
 QY 121 TPAPPEISAMVLTMMKEAFYLGKVTTHAVVYVPAFMDAROKTACGTAGANVWKI 180
 DB 121 tpaapeisamvltmmkeafylgkvtthavvypafmdaroktagtagnvwmki 180
 QY 181 INEPTAAAIAYGLDKREGERNILVFDLGGSTFDVSLTTDNGVFVYVATNGDTHLGEDF 240
 DB 181 ineptaaaiaygldkregernilvfdlggstfdvslttngvfvyvatngdthlggedf 240
 QY 241 DQVWEHETIKYKKKTKDVKRRAVOKTRREVEKAKALSSOQARLETSEFEGEDF 300
 DB 241 dqvwehetikykkktdvkravoktrrevkakalssogaretsefegedf 300
 QY 301 SETLFRKFEELNMDLFRSTMKPVQKVLNEDSLKSDIDEIVLVGSTRIPKIQOLVKEF 360
 DB 301 setlfrakfeelnmdlfrstmkpvqkvlndsdlksdideivlvgstripkiglvkef 360
 QY 361 FNGKPSHSIGINPDEAVAYGAAGVSGDPDGDVLLDNGPRLTGETGQGWTKRIIP 420
 DB 361 fngkpsysiginpdeavaygaagvsgdpgdvlldngprltgetgqgwtkriip 420
 QY 421 RNTVVPYTKSSQIFSTASDNPYVTVTKVYEGEERPLJNDMLGLTFDLTGIPAPRGVPIE 480
 DB 421 rntvvpkyksqifstasdnpvtytvtkvyegerpljndmlglftdlgtgipprgvpjie 480
 QY 481 VFEEDVNGILIRVTAEDKGGKGNKKITITNDONRLPELTERVNDAEKFAEDKKKER 540
 DB 481 vfeedvngilrvtaedkggknkkititndonrlpeltervndaekfaedkkker 540
 QY 541 IDTRNELESYAVSLKNOIGDKEKLGKLSSEDEKTEIMKAVEEKIEWLESHODADIEDPKA 600
 DB 541 idtrnelesyavslknoigdekkgklsessedeketiemkaveekiewleshodadiedfka 600

QY 601 KKELEEVOPITISKLYGSAGPPPTGEEDTAEL 633
 DB 601 kkeleevopitisklygsagppptgeedtael 633

RESULT 2

ID AAY90693 standard; Protein: 639 AA.

AC AAY90693:

DT 29-AUG-2000 (first entry)

DE 639 residue human immunoglobulin heavy binding protein, BIP(78KD).

KW Immunoglobulin heavy chain binding protein: BIP(78KD); chondrocyte;
 KW autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.

OS Homo sapiens.

PN WO200021995-A1.

PD 20-APR-2000.

PF 08-OCT-1999; 99MO-GB03316.

PR 09-OCT-1998; 98GB-0022115.

PA (UNLO) KINGS COLLEGE LONDON.

PI Panayl GS, Corriganll VM, Bodman-Smith MD, Fife MS, Lanchbury JS;

DR WPI: 2000-317942/27.

DR N-PSDB: AAA30792.

PT New human immunoglobulin heavy chain binding protein and encoding

PT polynucleotide, useful for diagnosis and treatment of rheumatoid

PT arthritis

PS Claim 3; Page 23; 53pp; English.

CC This sequence represents a human immunoglobulin heavy chain binding

CC protein, BIP(78KD), having a 633 amino acid sequence. The invention

CC also encompasses a BIP(78KD) protein of 639 amino acids (Y90694).

CC The cDNA encoding BIP(78KD), also referred to as p78 in the

CC specification, was isolated from human chondrocytes (the specialised

CC cells of articular cartilage) and human chondrosarcoma cell lines. The

CC BIP(78KD) cDNA of this invention contains a number of differences

CC compared with the published sequence (Genbank accession number X87949),

CC and has therefore been deposited with Genbank with the accession

CC number AF188611). These differences comprise 6 single nucleotide

CC substitutions and a codon insertion, and result in three amino acid

CC substitutions and an arginine insertion at position 834-836 of the

CC protein. The BIP(78KD) proteins react with antibodies present in the

CC serum of Rheumatoid arthritis patients, and is therefore a putative

CC autoantigen for this autoimmune disease. BIP(78KD) is also able to

CC selectively proliferate synovial T-cells from patients with rheumatoid

CC arthritis. BIP(78KD) or peptides derived from the protein are useful as

CC reagents to indicate the presence of rheumatoid arthritis, and can be

CC used in prognostic or diagnostic tests of body fluids for rheumatoid

CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western

CC blotting. The protein or the cDNA encoding it can also be used to test

CC for rheumatoid arthritis by detecting antibodies to the protein.

CC BIP(78KD), its peptides and polynucleotides are also useful

CC therapeutically.

XX Sequence 639 AA:

Query Match 100.0%; Score 3225; DB 21; Length 639;
 Best Local Similarity 100.0%; Pred. No. 2.3e-208;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:57 ; Search time 18.76 seconds
(without alignments)
1248.871 Million cell updates/sec

Title: US-09-806-955-1

Sequence: 1 MEEDKKEVDVGTVGIDLT.....ACPPPTGEEDTAELHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

1 number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SWISSProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	3216	98.3	654	1 GR78_HUMAN	P11021 homo sapien
2	3209	98.0	654	1 GR78_MOUSE	P07823 mesocricetu
3	3205	97.9	655	1 GR78_MOUSE	P20029 mus musculus
4	3200	97.8	654	1 GR78_RAT	P06761 ratu mus norv
5	3168	96.8	652	1 GR78_CHICK	Q90593 gallus gall
6	3069	93.8	658	1 GR78_XENLA	Q91883 xenopus lae
7	2722.5	83.2	667	1 GR78_APLCA	Q16956 aplysia cal
8	2708	82.7	656	1 HS7C_DROME	P23440 drosophila
9	2594.5	79.3	661	1 HS7C_CAEEL	Q03684 nicotiana t
10	2356.5	72.0	667	1 BIP3_TOBAC	Q03685 nicotiana t
11	2347.5	71.7	668	1 BIP3_TOBAC	Q03685 nicotiana t
12	2333.5	71.3	663	1 BIP3_MAIZE	O24581 zea mays (m
13	2328.5	71.1	666	1 BIP3_LYCES	P45118 lycopersico
14	2328.5	71.1	668	1 BIP3_SPIOL	Q42434 spinacia ol
15	2322.5	71.0	663	1 BIP3_MAIZE	P24067 zea mays (m
16	2301	70.3	669	1 BIP3_ARATH	Q39433 arabidopsis
17	2299	70.2	668	1 BIP3_ARATH	P76695 neurospora
18	2227.5	68.1	662	1 GR78_NEUCR	P36604 schizosach
19	2216	67.1	663	1 GR78_SCHPO	Q39170 yarrowia li
20	2195	67.1	670	1 GR78_TARLI	P16474 saccharomyc
21	2174	66.4	682	1 GR78_YEAST	Q05866 plasmodium
22	2100	64.2	655	1 GR78_PLAFO	P10591 saccharomyc
23	2085.5	63.7	641	1 HS71_YEAST	P11142 homo sapien
24	2083	63.6	646	1 HS7C_HUMAN	P08109 mus musculu
25	2079	63.5	650	1 HS7C_MOUSE	P19120 bos taurus
26	2078	63.5	644	1 HS71_BOVIN	P53421 pichia angu
27	2076.5	63.5	640	1 HS71_PICAN	P36415 dictyostell
28	2074.5	63.4	653	1 HS71_CANAL	P41797 candida alb
29	2074	63.4	646	1 HS72_CRIGR	P13378 cricetus
30	2072	63.3	638	1 HS72_YEAST	P10592 saccharomyc
31	2065.5	63.1	679	1 GR78_KLUFA	P22010 kluyveromyc
32	2065.5	63.1	648	1 HS71_PUCGR	Q01877 puccinia gr

34	2064	63.1	641	1 HS7A_YEAST	P22202 saccharomyc
35	2063.5	63.0	649	1 HS71_YEAST	P09435 saccharomyc
36	2060.5	63.0	646	1 HS70_NEUCR	O01233 neurospora
37	2060	62.9	643	1 HS71_SCHPO	O10265 schizosach
38	2059	62.9	645	1 HS70_SOYBN	P26413 glycine max
39	2053.5	62.7	651	1 HS7C_PETHY	P09189 petunia hyb
40	2052	62.7	652	1 HS7D_MANSE	P09189 petunia sex
41	2049	62.6	651	1 HS70_ONCMY	P08108 oncorhynch
42	2046	62.5	641	1 HS72_BOVIN	O27965 bos taurus
43	2045	62.5	651	1 HS71_ARATH	P22955 arabidopsis
44	2043.5	62.4	639	1 HS72_HUMAN	P34652 homo sapien
45	2039.5	62.3	641	1 HS71_HUMAN	P08107 homo sapien

ALIGNMENTS

RESULT 1
GR78_HUMAN STANDARD; PRT; 654 AA.
ID GR78_HUMAN
AC P11021; Q9NPP1;
DT 01-JUL-1989 (Rel. 11, Created)
DF 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN
DE HEAVY CHAIN BINDING PROTEIN) (BIP) (ENDOPLASMIC RETICULUM LUMENAL CA2+
DE BINDING PROTEIN GRP78).
GN HSP45 OR GRP78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88283347; PubMed=2840249;
RA Ting J., Lee A.S.;
RT Human gene encoding the 78,000-dalton glucose-regulated protein and
RT its pseudogene: structure, conservation, and regulation.*;
RL DNA 7:275-286(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Cervix carcinoma;
RA Chao C.C.K.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Fibroblast;
RA Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;
RT Grp78 is involved in the quality control of the LDL-receptor.*;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Bernudez-Fajardo A., Ilewelyn D.H., Campbell A.K., Errington R.R.;
RT Sequence differences between human grp78/BIP isolated from HeLa cells
RT and previously reported human sequences.*;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-25 FROM N.A.
RX MEDLINE=93126087; PubMed=1480470;
RA Chao C.C.K., Lin-Chao S.;
RT A direct-repeat sequence of the human BIP gene is required for
RT A23187-mediated inducibility and an inducible nuclear factor
RT binding.*;
RL Nucleic Acids Res. 20:6481-6485(1992).
RN [6]
RP SEQUENCE OF 22-38.
RX TISSUE=Breast carcinoma;
RA MEDLINE=97295304; PubMed=9150946;
RA Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.,
RA Simpson R.J., Dorow D.S.;
RT Two-dimensional electrophoretic analysis of human breast carcinoma
RT proteins: mapping of proteins that bind to the SH3 domain of mixed
RT lineage kinase MLK2.*;

```

RN Electrophoresis 18:588-598(1997).
RL (7)
RP SEQUENCE OF 19-40.
RC TISSUE-COLON CARCINOMA;
RX MEDLINE-97295306; PubMed-9150948;
RA J.H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
RT "A two-dimensional gel database of human colon carcinoma proteins.";
RL Electrophoresis 18:605-613(1997).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19645; AAA52614.1;
DR EMBL; X87949; CAA61201.1;
DR EMBL; AJ271729; CAB71335.1;
DR EMBL; AF216292; AAF42836.1;
DR EMBL; X59969; CAA42595.1;
DR PIR; A29821; A29821.
DR HSP; P19120; IATR.
DR SWISS-2DPAGE; P11021; HUMAN.
DR MIM; 138120;
DR InterPro; IPR000886; ER_target.
DR InterPro; IPR001023; HSP70.
DR Pfam; PF00012; HSP70.1.
DR PRINTS; PR00301; HEATSHOCK70.
DR PROSITE; PS00014; ER_TARGET.1.
DR PROSITE; PS00329; HSP70_1;
DR PROSITE; PS00329; HSP70_2;
DR PROSITE; PS01036; HSP70_3;
DR ATP-binding; Endoplasmic reticulum; signal.
KW SIGNAL.
FT CHAIN 1 18
FT SITE 651 654 78 KDA GLUCOSE-REGULATED PROTEIN.
FT CONFLICT 297 297 MISSING (IN REF. 1 AND 2).
FT CONFLICT 418 418 D -> H (IN REF. 1 AND 2).
FT CONFLICT 439 439 R -> S (IN REF. 1 AND 2).
FT CONFLICT 447 447 K -> N (IN REF. 1 AND 2).
SQ SEQUENCE 654 AA; 72333 MW; 5987D8D5B32A00 CRC64;

```

```

Query Match 98.3%; Score 3216; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 1.5e-151;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 302 ETLRAFEELNMDLRSTKPKVQKYLESDLKSDIDEIYLVGSTRIRKIQLVKEFF 361
DB 320 ETLTRAFEEELNMDLFRSTKPKVQKYLESDLKSDIDEIYLVGSTRIRKIQLVKEFF 379
QY 362 NGKEDPSGKINPDEAVAVGAAGVAGVSGDQDGDGVLLDVCPPLTGLIETGVVATKILPR 421
DB 380 NGKEDPSGKINPDEAVAVGAAGVAGVSGDQDGDGVLLDVCPPLTGLIETGVVATKILPR 439
QY 422 NTVPPTKSDIETASDNOPTVTKYEGEERLTNDNLSTFPLDTGPPAPRCPVPIEV 481
DB 440 NTVPPTKSDIETASDNOPTVTKYEGEERLTNDNLSTFPLDTGPPAPRCPVPIEV 499
QY 482 TFEIDVNGILRVTAEDKGTGKNNKITTNDQNRLLPEEIERVNDAEKFAEEDKKLKERI 541
DB 500 TFEIDVNGILRVTAEDKGTGKNNKITTNDQNRLLPEEIERVNDAEKFAEEDKKLKERI 559
QY 542 DTRHELESYVSLKNOIGKKEKLGKLSSEDEKTEKAVEKIEMLSEHODADIEDPKAK 601
DB 560 DTRHELESYVSLKNOIGKKEKLGKLSSEDEKTEKAVEKIEMLSEHODADIEDPKAK 619
QY 602 KKELEETVQPTISKLYGSAGAPPTGEEDPAE 612
DB 620 KKELEETVQPTISKLYGSAGAPPTGEEDPAE 650

```

```

RESULT 2
GR78_MESAU STANDARD; PRT; 654 AA.
AC P07823;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last annotation update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 78 KDA GLUCOSE-REGULATED PROTEIN PRECURSOR (GRP 78) (IMMUNOGLOBULIN
DE HEAVY CHAIN BINDING PROTEIN) (BIP).
GN HSP45 OR GRP78.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-67305586; PubMed-3623104;
RA Tling J., Wooden S.K., Kriz R., Kelleher K., Kaufman R.J., Lee A.S.;
RT "The nucleotide sequence encoding the hamster 78-kDa
RT glucose-regulated protein (GRP78) and its conservation between
RT hamster and rat.";
RL Gene 55:147-152(1987).
CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN FACILITATING THE ASSEMBLY OF
CC MULTIMERIC PROTEIN COMPLEXES INSIDE THE ER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17169; AAA51448.1;
DR PIR; A27414; A27414.
DR HSP; P19120; IATR.
DR InterPro; IPR000886; ER_target.
DR InterPro; IPR001023; HSP70.
DR Pfam; PF00012; HSP70.1.
DR PRINTS; PR00301; HEATSHOCK70.
DR PROSITE; PS00014; ER_TARGET.1.
DR PROSITE; PS00329; HSP70_1;
DR PROSITE; PS00329; HSP70_2;
DR PROSITE; PS01036; HSP70_3;
KW ATP-binding; Endoplasmic reticulum; signal.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 30.91 Seconds

(without alignments)
1574.751 Million cell updates/sec

Title: us-09-806-955-1

Perfect score: 3273
Sequence: 1 MEDKKEDVGTGVGIDCTT.....AGPPPGEDTALNNNNH 639Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Minimum number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_68: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	2 A27414	dnak-type molecule
2	3201	97.8	653	2 A37048	dnak-type molecule
3	3200	97.8	654	1 HHRGB	dnak-type molecule
4	3182.5	97.2	653	2 A29821	dnak-type molecule
5	3168	96.8	652	2 I50242	dnak-type molecule
6	2723.5	83.2	667	2 S24782	dnak-type molecule
7	2719.5	83.1	667	2 D44261	dnak-type molecule
8	2710	82.8	656	2 JN0666	dnak-type molecule
9	2618.5	80.0	661	2 T15513	heat shock 70k pro
10	2594.5	79.3	661	2 A32475	dnak-type molecule
11	2559	78.2	657	2 T34037	heat shock 70k pro
12	2356.5	72.0	667	2 S21879	dnak-type molecule
13	2347.5	71.7	668	2 S21880	dnak-type molecule
14	2333.5	71.3	663	2 T04080	dnak-type molecule
15	2330	71.2	668	2 T46574	dnak-type molecule
16	2323.5	71.0	663	2 T04078	dnak-type molecule
17	2305.5	70.4	663	2 T03581	dnak-type molecule
18	2298	70.2	668	2 S71171	dnak-type molecule
19	2285	69.8	672	2 T43723	dnak-type molecule
20	2278	69.6	672	2 T43716	dnak-type molecule
21	2261	69.1	642	2 T45298	dnak-type molecule
22	2240.5	68.5	664	2 T06598	dnak-type molecule
23	2237	68.3	655	2 H86222	hypothetical prote
24	2231	68.2	658	2 S38890	glucose-regulated
25	2227.5	68.1	662	2 T50464	78 kd glucose regu
26	2216	67.7	663	2 T38155	dnak-type molecule
27	2214.5	67.7	608	2 T05741	dnak-type molecule
28	2205	67.4	666	2 T06358	dnak-type molecule
29	2174	66.4	682	1 HBBYK2	dnak-type molecule

30	2156	65.9	663	2 S20877	dnak-type molecule
31	2108	64.4	651	2 JC7132	heat shock protein
32	2100	64.2	655	2 A48468	dnak-type molecule
33	2083	63.6	646	2 S07197	dnak-type molecule
34	2083	63.6	646	2 JC4853	dnak-type molecule
35	2083	63.6	646	2 A27077	dnak-type molecule
36	2080	63.6	641	2 PC7036	heat shock protein
37	2079	63.5	647	2 T41121	heat shock protein
38	2079	63.5	650	2 S11456	dnak-type molecule
39	2078	63.5	645	2 S41372	dnak-type molecule
40	2077	63.5	646	2 A45935	dnak-type molecule
41	2076.5	63.4	640	2 S37394	dnak-type molecule
42	2074.5	63.4	648	2 S53498	dnak-type molecule
43	2074.5	63.4	656	2 S51712	dnak-type molecule
44	2074	63.4	646	2 A35922	dnak-type molecule
45	2072	63.3	639	2 S20139	dnak-type molecule

ALIGNMENTS

RESULT 1
A27414
dnak-type molecular chaperone GRP78 precursor - Chinese hamster
N:Alternate names: glucose-regulated 78k protein; Ig heavy chain-binding protein
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
C:Accession: A27414
R:Ring, J.; Wooden, S.K.; Kitz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.
Gene 55, 147-152, 1987
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei
A:Reference number: A27414; MUID:87305586
A:Accession: A27414
A:Molecule type: mRNA
A:Residues: 1-654 <TIN>
A:Cross-references: GH:M1769; NID:q191090; PIDN:AA51448.1; PID:q304510
C:Genetics:
A:Gene: GRP78
C:Function:
A:Description: Involved in protein folding and assembling/disassembling of protein co
C:Superfamily: heat shock protein 70
C:Keywords: ATP; heterotetramer; immunoglobulin; molecular chaperone
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-554/Product: Ig heavy chain binding protein #status predicted <MAT>
F:551-654/Region: endoplasmic reticulum retention signal

Query Match 98.0%; Score 3209; DB 2; Length 654;
Best local similarity 99.7%; Pred. No. 7.3e-155;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 EEDKKEDVGTGVGIDCTTSCVGFKNRGREIITANOGRRITPSVATPDEERLIGDA 61
DB 20 EEDKKEDVGTGVGIDCTTSCVGFKNRGREIITANOGRRITPSVATPDEERLIGDA 79
QY 62 ANKQLTSPNENTVPDAKLIGRTWNDSVODIKFLPFKVEKKTPYIOVDIGGQRTK 121
DB 80 ANKQLTSPNENTVPDAKLIGRTWNDSVODIKFLPFKVEKKTPYIOVDIGGQRTK 139
QY 122 FAPPEISAMVLTAKMETAEVIGKVVHAYTVAYVNDORATDAGTIGLWVRI 181
DB 140 FAPPEISAMVLTAKMETAEVIGKVVHAYTVAYVNDORATDAGTIGLWVRI 199
QY 182 NEPTAAAIAYGLDKREGSKNLTVPDLGGTFVSLTIDNGVEVVATNGDTHLGEQPD 241
DB 200 NEPTAAAIAYGLDKREGSKNLTVPDLGGTFVSLTIDNGVEVVATNGDTHLGEQPD 259
QY 242 QRYMEHFTLYKKTKGKVRKDNRAVOKLREVKAKRAISQHORIEESPEGEQDS 301
DB 260 QRYMEHFTLYKKTKGKVRKDNRAVOKLREVKAKRAISQHORIEESPEGEQDS 319
QY 302 ETLTRAKFEELNMDLFRSTMKPVOKVLEDSDLKSSDIDETIVLGSSFTRIQQLVKEFF 361

```

Db 320 ETLTRAFEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 379
Qy 362 NGKPSRGINPDEAVAYGAAGVAGVLSGDQDQDGLVLLDVCPPLTGLIETVGGVMTKILPR 421
Db 380 NGKPSRGINPDEAVAYGAAGVAGVLSGDQDQDGLVLLDVCPPLTGLIETVGGVMTKILPR 439
Qy 422 NTVPPTKKSQIFSTASDNQPTVIKYEGEERPLTKDNHLLGTFDLTGTPPAPRGVPOIEV 481
Db 440 NTVPPTKKSQIFSTASDNQPTVIKYEGEERPLTKDNHLLGTFDLTGTPPAPRGVPOIEV 499
Qy 482 TFEIDVNGILRVTAEDKGTGKGNKKITITNDQNRLLPEELERNVDAEKFAEDKKLKERT 541
Db 500 TFEIDVNGILRVTAEDKGTGKGNKKITITNDQNRLLPEELERNVDAEKFAEDKKLKERT 559
Qy 542 DTRNELSVAYSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLSHODADIEDPKAK 601
Db 560 DTRNELSVAYSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLSHODADIEDPKAK 619
Qy 602 KKELEIVQPIISKLYGSAGPPTGGEEDTAE 632
Db 620 KKELEIVQPIISKLYGSAGPPTGGEEDTAE 650

RESULT 2
A37048
dnak-type molecular chaperone grp78 precursor - mouse
M:Alternate names: BIP: glucose-regulated 78k protein; Ig heavy chain-binding protein
C:Species: Mus musculus (house mouse)
C:Date: 15-Jan-1993 #sequence-revision 15-Jan-1993 #text-change 20-Aug-1999
C:Accession: A37048; B37048; J00094; I49137; A31934
R:Kotsumu, Y.; Northington, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.
J. Cell Sci. Suppl. 11, 115-137, 1989
A:Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulated
A:Reference number: A37048; M0ID:90130686
A:Accession: A37048
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-655 <K21>
A:Accession: B37048
A:Molecule type: protein
A:Residues: 20-41, 'X', 43-46 <K22>
R:Parfett, C.L.J.; Hofbauer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.
Gene 82, 291-303, 1989
A:Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero
A:Reference number: J00094; M0ID:90060818
A:Accession: J00094
A:Molecule type: mRNA
A:Residues: 488-590, 'G', 592-595, 'E', 597-655 <PAR>
C:Cross-references: GB:M430779; NID:9193644; PIDN:AAA37742.1; PID:9193645
R:Tillman, J.B.; More, P.L.; Walford, R.L.; Spindler, S.R.
Gene 150, 225-229, 1995
A:Title: Structure and regulation of the mouse GRP78 (BIP) promoter by glucose and calci
A:Reference number: I49137; M0ID:95331621
A:Accession: I49137
A:Status: translation not shown; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-28 <RES>
R:Haas, I.G.; Meo, T.
Proc. Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988
A:Title: cDNA cloning of the immunoglobulin heavy chain binding protein.
A:Reference number: A31934; M0ID:88176922
A:Accession: A31934
A:Molecule type: mRNA
A:Residues: 514-589, 'D', 591-595, 'E', 597-655 <HAA>
A:Cross-references: GB:M19351; NID:9192231; PIDN:AAA37315.1; PID:9387113
C:Genetics:
A:Gene: grp78
C:Function:
A:Description: Involved in protein folding and assembling/dissassembling of protein comp
A:Superfamily: heat shock protein 70
C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon
F:1-19/Domain: signal sequence #status predicted <SIG>

```

```

F:20-655/Product: dnak-type molecular chaperone grp78 #status experimental <MAT>
F:652-655/Region: endoplasmic reticulum retention signal

Query Match          97.8%; Score 3201; DB 2; Length 655;
Best Local Similarity 99.4%; Pred. No. 1,8e-154;
Matches 627; Conservatve 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEDKKEDVGTGVGIDGCTTSCYGVFKNGRVEIIANDQNRITPSYAFPEGERLLGDA 61
Db 21 EEDKKEDVGTGVGIDGCTTSCYGVFKNGRVEIIANDQNRITPSYAFPEGERLLGDA 80
Qy 62 AKNOLTSNPEVTFDAKRLIGFTWNPVSQODIKFLPKRYVEKKRPYIQVDIGGGCTT 121
Db 81 AKNOLTSNPEVTFDAKRLIGFTWNPVSQODIKFLPKRYVEKKRPYIQVDIGGGCTT 140
Qy 122 PAPERISAMVLTWKKEETAEYLGKVTAAVTPVPFNDARQATKAGTITAGLNVKRII 181
Db 141 PAPERISAMVLTWKKEETAEYLGKVTAAVTPVPFNDARQATKAGTITAGLNVKRII 200
Qy 182 NEPTAAAIAYGLDKRDEKKNILVFDLGCGTTPVSLITDNGVEFVATNGDTLHGSDPD 241
Db 201 NEPTAAAIAYGLDKRDEKKNILVFDLGCGTTPVSLITDNGVEFVATNGDTLHGSDPD 260
Qy 242 QRVNHEFIKLYKKTKGDKVAKNRANVQKLRREYKAKRALSSQOARIEISFEGEDFS 301
Db 261 QRVNHEFIKLYKKTKGDKVAKNRANVQKLRREYKAKRALSSQOARIEISFEGEDFS 320
Qy 302 ETLTRAFEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 361
Db 321 ETLTRAFEEELNMDLFRSTMKPVOKVLESDLSKSDIDEIVLVGGSTRIPKIQOLVKEFF 380
Qy 362 NGKPSRGINPDEAVAYGAAGVAGVLSGDQDQDGLVLLDVCPPLTGLIETVGGVMTKILPR 421
Db 381 NGKPSRGINPDEAVAYGAAGVAGVLSGDQDQDGLVLLDVCPPLTGLIETVGGVMTKILPR 440
Qy 422 NTVPPTKKSQIFSTASDNQPTVIKYEGEERPLTKDNHLLGTFDLTGTPPAPRGVPOIEV 481
Db 441 NTVPPTKKSQIFSTASDNQPTVIKYEGEERPLTKDNHLLGTFDLTGTPPAPRGVPOIEV 500
Qy 482 TFEIDVNGILRVTAEDKGTGKGNKKITITNDQNRLLPEELERNVDAEKFAEDKKLKERT 541
Db 501 TFEIDVNGILRVTAEDKGTGKGNKKITITNDQNRLLPEELERNVDAEKFAEDKKLKERT 560
Qy 542 DTRNELSVAYSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLSHODADIEDPKAK 601
Db 561 DTRNELSVAYSLKNOIGDKKLGKLSSEDEKTEMEKAVEEKIEMLSHODADIEDPKAK 620
Qy 602 KKELEIVQPIISKLYGSAGPPTGGEEDTAE 632
Db 621 KKELEIVQPIISKLYGSAGPPTGGEEDTAE 651

RESULT 3
HHRITGB
dnak-type molecular chaperone precursor - rat
M:Alternate names: BIP: glucose-regulated 78k protein; Ig heavy chain-binding protein
N:Contains: steroidogenesis-activator polypeptide
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 22-Jun-1999
C:Accession: A23948; A60134; A26257
R:Munro, S.; Pelham, H.R.B.
Cell 46, 291-300, 1986
A:Title: An Hsp70-like protein in the ER: Identically with the 78 Kd glucose-regulated p
A:Reference number: A23948; M0ID:86245075
A:Accession: A23948
A:Molecule type: mRNA
A:Residues: 1-654 <MUN>
A:Cross-references: GB:M14050; NID:9203150; PIDN:AAA40817.1; PID:9203151
R:Pedersen, R.C.; Browne, A.C.
Science 236, 188-190, 1987
A:Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor.
A:Reference number: A60134; M0ID:87177581

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 43 Seconds

(without alignments)
1100.764 Million cell updates/sec

Title: us-09-806-955-1

Perfect score: 3273
Sequence: 1 MEEDKKEGVGVVIGIDGTT.....AGPPTEEDTAEILHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Minimum number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101:*

1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	21	AAV90693
2	3225	98.5	633	21	AAV90694
3	3209	98.0	654	15	AAAB3075
4	3182.5	97.2	653	22	AAAB2533
5	3171	96.9	656	14	AAAB4934
6	3168	96.8	666	15	AAAB3076
7	3289	70.2	668	21	AAAB1382
8	3259	70.2	691	21	AAAB1381
9	3258.5	69.0	658	21	AAAB39285
10	3258.5	69.0	658	21	AAAB39284
11	3237	68.3	655	21	AAAB30912

12	2174	66.4	682	22	AAAB70749
13	2125.5	64.9	687	22	AAAB70875
14	2083	63.6	646	19	AAAB4364
15	2083	63.6	646	20	AAAB17407
16	2083	63.6	646	20	AAAB17408
17	2083	63.6	646	21	AAAB23649
18	2083	63.6	646	21	AAAB23651
19	2083	63.6	646	22	AAAB2535
20	2083	63.6	890	21	AAAB22938
21	2074.5	63.4	656	18	AAAB01638
22	2064	63.1	645	11	AAAB03927
23	2061	62.4	647	11	AAAB03928
24	2039.5	62.3	624	11	AAAB23252
25	2039.5	62.3	641	19	AAAB43439
26	2039.5	62.3	641	21	AAAB23652
27	2039.5	62.3	641	22	AAAB2534
28	2036.5	62.2	646	21	AAAB3604
29	2033.5	62.1	633	14	AAAB3002
30	2032.5	62.1	647	20	AAAB1380
31	2032.5	62.1	647	22	AAAB49039
32	2032	62.1	634	11	AAAB03930
33	2021	61.7	640	11	AAAB23653
34	2017.5	61.6	640	11	AAAB03929
35	2014	61.5	642	11	AAAB23650
36	2013	61.5	640	18	AAAB10065
37	2013	61.5	640	21	AAAB8408
38	2013	61.5	677	11	AAAB09418
39	2008	61.4	641	14	AAAB3004
40	2004.5	60.9	669	11	AAAB03925
41	1994.5	60.9	676	9	AAAB0088
42	1974	60.3	665	21	AAAB58386
43	1970	60.2	646	11	AAAB03931
44	1968.5	60.1	657	11	AAAB03926
45	1963.5	60.0	623	21	AAAB53605

ALIGNMENTS

RESULT 1

AAV90693

ID AAV90693 standard; Protein: 639 AA.

AC AAV90693;

DT 29-AUG-2000 (first entry)

XX

DE 639 residue human immunoglobulin heavy binding protein, B1P(78KD).

KW Immunoglobulin heavy chain binding protein; B1P(78KD); chondrocyte; autoantigen; rheumatoid arthritis; antiarthritis; antirheumatic; p78.

OS Homo sapiens.

XX

PN WO200021995-A1.

XX

PD 20-APR-2000.

XX

PF 08-OCT-1999; 99MO-GB03316.

XX

PR 09-OCT-1998; 98GB-0022115.

XX

XX (UNLO) KINGS COLLEGE LONDON.

PA Penayl GS, Corriganl VM, Bodman-Smith MD, Fife MS, Lanchbury JS; WPI: 2000-317942/27.

XX

DR N-PSDB; AAAB30792.

XX

PT New human immunoglobulin heavy chain binding protein and encoding polynucleotide, useful for diagnosis and treatment of rheumatoid arthritis -

XX

PS Claim 3; Page 23; 53pp; English.

CC This sequence represents a human immunoglobulin heavy chain binding
CC protein, B1P(78KD), having a 639 amino acid sequence. The invention
CC also encompasses a B1P(78KD) protein of 633 amino acids (Y90694).
CC The cDNA encoding B1P(78KD), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC B1P(78KD) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949),
CC and has therefore been deposited with Genbank with the accession
CC number AF188611). These differences comprise 6 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The B1P(78KD) proteins react with antibodies present in the
CC serum of rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. B1P(78KD) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. B1P(78KD) or peptides derived from the protein are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC B1P(78KD), its peptides and polynucleotides are also useful
CC therapeutically.

XX Sequence 639 AA:

Query Match 100.0%; Score 3273; DB 21; Length 639;
Best Local Similarity 100.0%; Pred. No. 1.3e-209;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEDKEDVGVVGVGIDGTTSCVGFKNRGVEITIANDOGNITSTSYAFPEGERLIGD 60
DB 1 meedkdedvgvvgvvgidgttscvgyfknrgvveitlandognitstsyafpegerlrigd 60
QY 61 AAKNQLTSNPENTVFDAKRLIGRTWNPDSVODIKLPEFKVVEKTKPKYIOWDIGCGQTK 120
DB 61 aaknqltsnpentvfadkrligrtwndpsvoodiklpefkvvektkpkyyiowdigcgqtk 120
QY 121 TEAPEEISAWVLTQMKETAAVYLGKVTAAVYVYPAVFMDAQRQTKANGITAGLNWRI 180
DB 121 teapeeisaawvltqmketaaavylgkvtaavvyvpaufmdaqrqtkangitaglnwri 180
QY 181 INEPTAAIAVGLDKREGENKILVFPDLAGGTFPDSILITDNGVFVVAATNDTHLGEDF 240
DB 181 ineptaaiavglDKREGENKILVFPDLAGGTFPDSILITDNGVFVVAATNDTHLGEDF 240
QY 241 DORWMEHPIKLYKKKKTKDVKRKNRAVOKLRREVEKAKALSSQAAIEIESPEEGDP 300
DB 241 dorwmehpiklykkkktkdvkrknraavoklrrevekakalssqaaieiespeegdp 300
QY 301 SETLIRAKFEELNMDLFRSTMKPVQVLEDSDLKSDIDEIVLGVGSTRIRKIQOLVKEF 360
DB 301 setlirakfeelnmdlfrstmkpvqvleedsdlkksdideivlvgvstrirkioolvkef 360
QY 361 FNGEPEPSGIMPDVAVYGAAGVLSGDODGCVLLDNGVPLRLGLETVGQWTKRIIP 420
DB 361 fngpepsgimpdvavyygaagvlsdodgcvllldngvplrlgletvgqwtkriip 420
QY 421 RNTVVPTRKSSQIFSTASDNDPVTIVYVEGEERPLTKDNHILGTFDLTGIRPAPRGVPOIE 480
DB 421 rntvvptrkssqifstasdnndpvtivvygeerpltkdnhilgtfdltgirpapgvpjie 480
QY 481 VTFEIDNGIIRVTAEDKGTGNKNTITINDORRLPPEIERMVNDARKPAEDKXIKER 540
DB 481 vtfeldngiirvtaedkgtnknntitindorrlppeiermvndarkpaedkxiker 540
QY 541 IDTRNELESTAVSLKNDIGKERLGSSEDEKTEWEKAVEERIEWLESHODADIEDPKA 600
DB 541 idtrnelestavslkndigkerlgssedektewekaveeriewleshodadiedfka 600

QY 601 KKELEETVOPITISKLYGSAGPPPTGDEDTAEIHHHHHH 639
DB 601 kkeleeeivpilislygsagppptgdeetlaehhhhh 639

RESULT 2

AA990694
ID AA990694 standard; Protein: 633 AA.

AA990694:

29-AUG-2000 (first entry)

633 residue human immunoglobulin heavy binding protein, B1P(78KD).

Immunoglobulin heavy chain binding protein; B1P(78KD); chondrocyte;
autoantigen; rheumatoid arthritis; antiarthritic; antirheumatic; p78.

Homo sapiens.

MO200021995-AL.

20-APR-2000.

08-OCT-1999; 99MO-GH03316.

09-OCT-1998; 98GB-0022115.

(UNLO) KINGS COLLEGE LONDON.

Panayl GS, Corriganl VM, Bodman-Smith MD, Fife MS, Lanchbury JS;

WPI: 2000-317942/27.

New human immunoglobulin heavy chain binding protein and encoding
polynucleotide, useful for diagnosis and treatment of rheumatoid
arthritis -

Claim 3; Page 44-46; 53pp; English.

CC This sequence represents a human immunoglobulin heavy chain binding
CC protein, B1P(78KD), having a 633 amino acid sequence. The invention
CC also encompasses a B1P(78KD) protein of 639 amino acids (Y90693).
CC The cDNA encoding B1P(78KD), also referred to as p78 in the
CC specification, was isolated from human chondrocytes (the specialised
CC cells of articular cartilage) and human chondrosarcoma cell lines. The
CC B1P(78KD) cDNA of this invention contains a number of differences
CC compared with the published sequence (Genbank accession number X87949),
CC and has therefore been deposited with Genbank with the accession
CC number AF188611). These differences comprise 6 single nucleotide
CC substitutions and a codon insertion, and result in three amino acid
CC substitutions and an arginine insertion at position 834-836 of the
CC protein. The B1P(78KD) proteins react with antibodies present in the
CC serum of rheumatoid arthritis patients, and is therefore a putative
CC autoantigen for this autoimmune disease. B1P(78KD) is also able to
CC selectively proliferate synovial T-cells from patients with rheumatoid
CC arthritis. B1P(78KD) or peptides derived from the protein are useful as
CC reagents to indicate the presence of rheumatoid arthritis, and can be
CC used in prognostic or diagnostic tests of body fluids for rheumatoid
CC arthritis by ELISA (enzyme linked immunosorbent assay) or Western
CC blotting. The protein or the cDNA encoding it can also be used to test
CC for rheumatoid arthritis by detecting antibodies to the protein.
CC B1P(78KD), its peptides and polynucleotides are also useful
CC therapeutically.

XX Sequence 633 AA:

Query Match 98.5%; Score 3225; DB 21; Length 633;
Best Local Similarity 100.0%; Pred. No. 1.9e-206;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:37 Search time 50.77 Seconds
(without alignments)

1841.009 Million cell updates/sec

Title: US-09-806-955-1
Sequence: 1 MEDKKEDVGVVGDIDLT.....AGPPTGEEDTAEIHHNNH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Minimum number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3273	100.0	639	09UK02	09UK02 homo sapien
2	3216	98.3	654	09NPF1	09NPF1 homo sapien
3	3199	97.7	655	11_09DC41	09DC41 mus musculu
4	2994	91.5	658	13_091688	091688 xenopus lae
5	2725	83.3	658	5_076180	076180 bombyx mori
6	2717	83.0	656	5_09VT03	09VT03 drosophila
7	2607.5	79.7	656	5_062568	062568 superilles d
8	2559	78.2	657	5_002085	002085 caenorhabdi
9	2534	77.4	651	5_024798	024798 echinococcu
10	2525.5	77.2	649	5_024895	024895 echinococcu
11	2460	75.2	648	5_045038	045038 schistosoma
12	2345.5	71.7	665	10_09M4E8	09M4E8 cucumis sat
13	2330	71.2	668	10_022639	022639 glycine max
14	2313.5	70.4	675	10_040924	040924 pseudotsuga
15	2303.5	70.4	663	10_024182	024182 oryza sativ
16	2292	70.0	668	10_09FSY7	09FSY7 corylus ave
17	2285	69.8	672	3_014453	014453 aspergillus
18	2278	69.6	672	3_013280	013280 aspergillus
19	2273	69.4	668	5_09U540	09U540 toxoplasma

20	2263	69.1	669	3_09C1C1	09C1C1 cronartium
21	2261	69.1	642	5_061001	061001 toxoplasma
22	2246.5	68.6	672	3_09WE3	09WE3 aspergillus
23	2243	68.5	659	10_041074	041074 phaeodactyl
24	2240.5	68.5	664	10_039830	039830 glycine max
25	2237	68.3	655	10_004022	004022 arabidopsis
26	2231	68.2	658	10_012752	012752 phytophthor
27	2229.5	68.1	701	5_024928	024928 elmeria ten
28	2214.5	67.7	608	10_040058	040058 hordeum vul
29	2205	67.4	656	10_039804	039804 glycine max
30	2188.5	66.9	655	3_P87028	P87028 pneumocysti
31	2160.5	66.0	665	3_09HG01	09HG01 pichia angu
32	2139.5	65.4	652	3_000053	000053 pneumocysti
33	2116.5	64.7	652	5_007615	007615 plasmodium
34	2108	64.4	651	3_09UVM1	09UMV1 rhizopus st
35	2093	63.9	659	5_09XZJ2	09XZJ2 crassostrea
36	2080	63.6	642	3_09UVM0	09UVM0 rhizopus st
37	2079	63.5	647	3_059855	059855 schizosacch
38	2076	63.4	661	5_077164	077164 entamoeba h
39	2074.5	63.4	648	10_040980	040980 pistum sativ
40	2069	63.2	646	13_073885	073885 gallus gall
41	2065.5	63.1	647	3_094106	094106 pneumocysti
42	2064.5	63.1	645	3_094106	094106 pneumocysti
43	2057	62.8	652	5_026924	026924 trypanosoma
44	2055.5	62.8	647	10_09M4E6	09M4E6 cucumis sat
45	2055.5	62.8	653	5_026733	026733 trypanosoma

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	639 AA.
ID	09UK02			
AC	09UK02:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	BIP PROTEIN (FRAGMENT).			
GN	HSPA5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI-TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=ARTICULAR CARTILAGE;			
RA	File M.S.:			
RT	*Bip.: an Antigen associated with Rheumatoid Arthritis.*;			
DR	EMBL: AF186811; AAF13605.1; -			
DR	HSP; P13120; HSC.			
DR	InterPro: IPR01023; HSP70.			
DR	Pfam: PF00012; HSP70.1.			
DR	PRINTS: PR00301; HEATSHOCK70.			
DR	PROSITE: PS00297; HSP70_1; 1.			
DR	PROSITE: PS00329; HSP70_2; 1.			
DR	PROSITE: PS01036; HSP70_3; 1.			
FT	NON_TER 639			
SO	SEQUENCE 639 AA; 70931 MW; 3421208FF128FD5D CRC64;			
Query Match	100.0%; Score 3273; DB 4; Length 639;			
Best Local Similarity	100.0%; Pred. No. 2,3e-158;			
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 MEDKKEDVGVVGDIDLTGYSYGVGRKNGVETIRANDGSRITPSVAFTEGEGRLGD 60			
DB	1 MEDKKEDVGVVGDIDLTGYSYGVGRKNGVETIRANDGSRITPSVAFTEGEGRLGD 60			
QY	AAAKNLTSPNTEVDFARRLIGRTWNPDSVVOODIKFLPFKVKVKTKPYIQVDIGGGQTK 120			
DB	AAAKNLTSPNTEVDFARRLIGRTWNPDSVVOODIKFLPFKVKVKTKPYIQVDIGGGQTK 120			

Oy	121	TFAPEEISAMVLTKKKEAEAYLGKKTVAHVYAVAFENDQORATDAGTIGLVNMI	180
Db	121	TFAPEEISAMVLTKKKEAEAYLGKKTVAHVYAVAFENDQORATDAGTIGLVNMI	180
Oy	181	INPFTAAAIAYGDKRGEKNILVFDLGGSTEDVSLITIDNGVEVVATNGDTHLGEDF	240
Db	181	INPFTAAAIAYGLDRREEKNILVFDLGGSTEDVSLITIDNGVEVVATNGDTHLGEDF	240
Oy	241	DQWMEHFIKLKTKKGTGDVKDNRAVQKLREVEKARALSSQHARIETESFYEGEDF	300
Db	241	DQWMEHFIKLKTKKGTGDVKDNRAVQKLREVEKARALSSQHARIETESFYEGEDF	300
Oy	301	SETLTAKKEELNMDFPSTMKPKQKVLESDLSKSIDELYVGSGSTRIPKIQQLUYEF	360
Db	301	SETLTAKKEELNMDFPSTMKPKQKVLESDLSKSIDELYVGSGSTRIPKIQQLUYEF	360
Oy	361	FNGEKPSRGINDENAVAGAAYAGVLSDDOTGDLVLLDVCPRLTGIEFGVGMTKLIP	420
Db	361	FNGEKPSRGINDENAVAGAAYAGVLSDDOTGDLVLLDVCPRLTGIEFGVGMTKLIP	420
Oy	421	RNVVTAKKSQIFSAASNOPRYTIKWYEGEERPLKDNNHLSGFEDLGIPTPARAGPOLE	480
Db	421	RNVVTAKKSQIFSAASNOPRYTIKWYEGEERPLKDNNHLSGFEDLGIPTPARAGPOLE	480
Oy	481	VFEFDVINGILRVTAEDKGCTGNKKTTITNDONRLTPRELIRMVADAEKFAEDKKLYER	540
Db	481	VFEFDVINGILRVTAEDKGCTGNKKTTITNDONRLTPRELIRMVADAEKFAEDKKLYER	540
Oy	541	IDVRNELESYAALKNOIGDKEKLGKLSSEKPEMKAWEKEIEMLESNQADIEDPKA	600
Db	541	IDVRNELESYAALKNOIGDKEKLGKLSSEKPEMKAWEKEIEMLESNQADIEDPKA	600
Oy	601	KKKELEEVPIPIISKLYGSAGPRPGEBDTAYELNNHHN 639	
Db	601	KKKELEEVPIPIISKLYGSAGPRPGEBDTAYELNNHHN 639	
<hr/>			
RESULT	2		
IONPF1	09NPFL	PRELIMINARY;	PRT: 654 AA.
AC	09NPFL:		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	GLUCOCSE-REGULATED PROTEIN (ENDOPLASMIC RETICULUM LUMINAL CA2+ BINDING PROTEIN GRP78).		
DE	HSPA5.		
OS	Homo sapiens (Human).		
CS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NBUI_TaxId=9606;		
RM	[1]		
RC	SEQUENCE FROM N.A.		
RP	TISSUE=EIBROBLASTS;		
RA	Hansen J.J., Nielsen M.N., Jorgensen M.M., Gregersen N., Bolund L.;		
RT	"Grp78 is involved in the quality control of the LDL-receptor.";		
RL	Submitted (JAN-2000) to the EMBL/genbank/DBJ databases.		
RL	[2]		
RC	SEQUENCE FROM N.A.		
RP	TISSUE=FIBROBLASTS;		
RA	Ghao C.;		
RL	Submitted (FEB-2000) to the EMBL/genbank/DBJ databases.		
RL	[3]		
RP	SEQUENCE FROM N.A.		
RA	Bermudez-Fajardo A., Llewellyn D.H., Campbell A.K., Errington R.R.;		
RT	"Sequence differences between human grp78/Bip isolated from HeLa cells		
RL	and previously reported human sequences.";		
RL	Submitted (DEC-1999) to the EMBL/genbank/DBJ databases.		
DR	EMBL: AJ271729; CAB/1335.1; -;		
DR	EMBL: AF216292; AAP42836.1; -;		
DR	Interpro: IPR000886; ER_target.		
DR	Interpro: IPR001023; HSP70.		

DR	Pfam: PF00012; HSP70_1.
DR	Prints: PR00301; HEATSHOCK70.
DR	Prosite: PS00014; ER_TARGET; UNKNOWN_1.
DR	Prosite: PS00297; HSP70_1; 1.
DR	Prosite: PS00329; HSP70_2; 1.
DR	Prosite: PS01036; HSP70_3; 1.
SQ	SEQUENCE 654 AA; 72332 MW; 59E7DB85BC32A00 CRC64;

Query Match	98.3%; Score 3216; DB 4; Length 654;
Best Local Similarity	100.0%; Pred. No. 1,9e-155;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

OY	2	EEDKKEVGVTVGIDGTTTSCVGFVKGRVETIANDGNITSTSYAFPEEGERLIGA	61
Dd	20	EEDKKEVGVTVGIDGTTTSCVGFVKGRVEIIANDGNITSTSYAFPEEGERLIGA	79
OY	62	AKNOLTSNPENTVPDAKKRLIGRTMNDPSVOODIKFLPKVYEKTKRYIVDYGCGOTKT	121
Dd	80	AKNOLTSNPENTVPDAKKRLIGRTMNDPSVOODIKFLPKVYEKTKRYIVDYGCGOTKT	139
OY	122	FAPBEISAVLVTKMETAEALVGKVVTAAYVTTPAFENDAQQTADAGTIAGLNWRII	181
Dd	140	FAPBEISAVLVTKMETAEALVGKVVTAAYVTTPAFENDAQQTADAGTIAGLNWRII	199
OY	182	NEPFAAAIAVGLDKRGEKNITLVFDLGSGTFPVSLTITDNGVFVANTNDTHLGEDPD	241
Dd	200	NEPFAAAIAVGLDKRGEKNITLVFDLGSGTFPVSLTITDNGVFVANTNDTHLGEDPD	259
OY	242	QRVMEHFILYKKKTKGDVARDNRNAVOKLREVEKAKRALSSOHARIETESFYEGEDS	301
Dd	260	QRVMEHFILYKKKTKGDVARDNRNAVOKLREVEKAKRALSSOHARIETESFYEGEDS	319
OY	302	ETTLRAFEELANDLFRSTMKPQVKYLESDDLKSDIDEIYLVGSTRIPRIQOLYKEFP	361
Dd	320	ETTLRAFEELANDLFRSTMKPQVKYLESDDLKSDIDEIYLVGSTRIRIKIQOLYKEFP	379
OY	362	NGKEPSSGIMPDENAVAIGAAGVAGLSGGDTGDI.VLLDVCP.LLTGLIEYGVWTKLIR	421
Dd	380	NGKEPSSGIMPDENAVAIGAAGVAGLSGGDTGDI.VLLDVCP.LLTGLIEYGVWTKLIR	439
OY	422	NTRYPTKKSQIFSTASDNQPTVIKYEEGERLTKDNHL.LTFPLTJGIPAPRGVPOLEY	481
Dd	440	NTRYPTKKSQIFSTASDNQPTVIKYEEGERLTKDNHL.LTFPLTJGIPAPRGVPOLEY	499
OY	482	TTELVDVGLIRVTAEDKGCTGNKKNTITTNDONRLTPEETEMRVADAEPEDKKLKERI	541
Dd	500	TTELVDVGLIRVTAEDKGCTGNKKNTITTNDONRLTPEETEMRVADAEPEDKKLKERI	559
OY	542	DTRELETSYAVALSKNOIGDKERLKGKLSEDEKTEKAVEEKIMLESHPDADIIDPFKAK	601
Dd	560	DTRELETSYAVALSKNOIGDKERLKGKLSEDEKTEKAVEEKIMLESHPDADIIDPFKAK	619
OY	602	KKELEETVOPITISKLYGSAGPPPTGEDTAE	632
Dd	620	KKELEETVOPITISKLYGSAGPPPTGEDTAE	650

RESULT 3	
Q9DC41	PRELIMINARY: PRT: 655 AA.
AC Q9DC41:	
DT 01-JUN-2001 (TrEMBLrel. 17, Created)	
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE HEAT SHOCK 70KD PROTEIN 5 (GLUCOSE-REGULATED PROTEIN, 78KD).	
GN HSPA5.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxId=10090;	
NN [1]	
RP SEQUENCE FROM N.A.	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:21:34 ; Search time 30.91 Seconds
(Without alignments)

1559.964 Million cell updates/sec

Title: US-09-806-955-2

Perfect score: 3225
Sequence: 1 MEEDKEDVGTVVIGDICTT.....SKLYGAGPPYGEEDTAEI 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

a1 number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	99.5	654	2 A27414	dnak-type molecule
2	3201	99.3	655	2 A37048	dnak-type molecule
3	3200	99.2	654	1 HHRTGB	dnak-type molecule
4	3182.5	98.7	653	2 A29821	dnak-type molecule
5	3168	98.2	652	2 I50242	dnak-type molecule
6	2722.5	84.4	667	2 S24782	dnak-type molecule
7	2719.5	84.3	667	2 D44261	dnak-type molecule
8	2710	84.0	656	2 JN0666	dnak-type molecule
9	2618.5	81.2	661	2 T15513	heat shock 70K pro
10	2594.5	80.4	661	2 A32475	dnak-type molecule
11	2555	79.2	657	2 T34037	heat shock 70K pro
12	2354.5	73.0	667	2 S21879	dnak-type molecule
13	2346.5	72.8	668	2 S21880	dnak-type molecule
14	2332.5	72.3	663	2 T04080	dnak-type molecule
15	2325	72.1	668	2 T46574	dnak-type molecule
16	2322.5	72.0	663	2 T04078	dnak-type molecule
17	2305.5	71.5	663	2 T03581	dnak-type molecule
18	2295	71.2	668	2 S71171	dnak-type molecule
19	2283	70.8	672	2 T43723	dnak-type molecule
20	2276	70.1	642	2 T43716	dnak-type molecule
21	2261	70.1	642	2 T45298	dnak-type molecule
22	2237	69.4	652	2 H86222	hypothetical prote
23	2234.5	69.3	664	2 T06598	dnak-type molecule
24	2231	69.2	658	2 S38890	dnak-type molecule
25	2227.5	69.1	662	2 T50464	glucose-regulated
26	2216	68.7	663	2 T38155	78 kD glucose regu
27	2214.5	68.7	608	2 T05741	dnak-type molecule
28	2200	68.2	666	2 T06358	dnak-type molecule
29	2173	67.4	682	1 HHBYK2	dnak-type molecule

ALIGNMENTS

RESULT	1	ALIGNMENTS
A27414	dnak-type molecular chaperone GRP78 precursor - Chinese hamster	
N:Alternate names: glucose-regulated 78K protein; Ig heavy chain-binding protein		
C:Species: Cricetus griseus (Chinese hamster)		
C:Date: 30-Jun-1988 #sequence-revision 30-Jun-1988 #text-change 20-Aug-1999		
C:Accession: A27414		
R:Ying, J.; Wooden, S.K.; Kriz, R.; Kelleher, K.; Kaufman, R.J.; Lee, A.S.		
Gene 55, 147-152, 1987		
A:Title: The nucleotide sequence encoding the hamster 78-kDa glucose-regulated protei		
A:Reference number: A27414; MUID:87305586		
A:Accession: A27414		
A:Molecule type: mRNA		
A:Residues: 1-654 <TIN>		
A:Cross-references: GB:M17169; NID:9191090; PIDN:AAA51448.1; PID:g304510		
C:Genetics:		
A:Gene: GRP78		
C:Function:		
A:Description: Involved in protein folding and assembling/dissassembling of protein co		
C:Superfamily: heat shock protein 70.		
C:Keywords: Arp; heterotrimer; immunoglobulin; molecular chaperone		
F:1-18/Domain: signal sequence #status predicted <SIG>		
F:19-654/Product: Ig heavy chain binding protein #status predicted <MAT>		
F:651-654/Region: endoplasmic reticulum retention signal		
Query Match	99.5%; Score 3209; DB 2; Length 654;	
Best Local Similarity	99.7%; Pred. No. 4e-155;	
Matches	629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	
OY	2 EEDKEDVGTVVIGDICTTSCYGVKNGRVEITINDGNRTTSYVAFPPGRLTGA 61	dnak-type molecule
DB	20 EEDKEDVGTVVIGDICTTSCYGVKNGRVEITINDGNRTTSYVAFPPGRLTGA 79	dnak-type molecule
OY	62 AKNQLTSPNPTVFDKRLRILGRTWNPVSQODIKFLPKVVEKTKRPVIGDGGGOTK 121	dnak-type molecule
DB	80 AKNQLTSPNPTVFDKRLRILGRTWNPVSQODIKFLPKVVEKTKRPVIGDGGGOTK 139	dnak-type molecule
OY	122 FAPBEISAMVLTAKKEAEVYLGKKTTHAVVTVVAVFENDAOQATKAGTLAGLVNRRIT 181	dnak-type molecule
DB	140 FAPBEISAMVLTAKKEAEVYLGKKTTHAVVTVVAVFENDAOQATKAGTLAGLVNRRIT 199	dnak-type molecule
OY	182 NEPTAAIAVGLDKRDEGENILVFDDGGCTFVSLTITDNGVFEVATNGDTHLGGDFD 241	dnak-type molecule
DB	200 NEPTAAIAVGLDKRDEGENILVFDDGGCTFVSLTITDNGVFEVATNGDTHLGGDFD 259	dnak-type molecule
OY	242 QRVNHEITIKLYKKKTGDKVAKNRVOKILREVEAKAKALSSQARLEISFEGDPS 301	dnak-type molecule
DB	260 QRVNHEITIKLYKKKTGDKVAKNRVOKILREVEAKAKALSSQARLEISFEGDPS 319	dnak-type molecule
OY	302 ETLTAAFEELINDLFRSTAKPVQKVEDSDLEKSDIDEIVLVGGSTRIPKIQOLVVEFF 361	dnak-type molecule

Db 320 ETLTRAKFEELNMDLFRSTMKPVQKVLIEDSDLKKSIDIELVVGSTRIPKIQOLVKEFF 379
Qy 362 NCKEPRSGINPDEAVANGAAGVAVSGSDODTDGLVLDVCPVLTLGIEVGVMTKLIPR 421
Db 380 NCKEPRSGINPDEAVANGAAGVAVSGSDODTDGLVLDVCPVLTLGIEVGVMTKLIPR 439
Qy 422 NTVPVPTKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFTDLTGIPAPRGVQIEV 481
Db 440 NTVPVPTKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFTDLTGIPAPRGVQIEV 499
Qy 482 TFEIDVNGILVLTAEADGKTGNKNTITNDQNLTPPEELERAVNDAEKFAEDDKLKXRI 541
Db 500 TFEIDVNGILVLTAEADGKTGNKNTITNDQNLTPPEELERAVNDAEKFAEDDKLKXRI 559
Qy 542 DTRNLESEYAVSLANKQIGDKELGKLSSEDKETMEKAVEKEIEMLESHQADIEDFKAK 601
Db 560 DTRNLESEYAVSLANKQIGDKELGKLSSEDKETMEKAVEKEIEMLESHQADIEDFKAK 619
Qy 602 KKELEIYQPIISKLKSGACPPPGEDDTAE 632
Db 620 KKELEIYQPIISKLKSGACPPPGEDDTAE 650

RESULT 2

A37048
dnaK-type molecular chaperone grp78 precursor - mouse
N:Alternate names: BiP; glucose-regulated 78k protein; Ig heavy chain-binding protein
C:Species: Mus musculus (house mouse)
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 20-Aug-1999
C:Accession: A37048; B37048; J00094; I49137; A31934
R:Kozaktsami, Y.; Norrlinson, K.; Press, E.; Slaughter, C.; Sambrook, J.; Gething, M.J.
A:Cell Sci. Suppl. 11, 115-137, 1989
A>Title: Identification of immunoglobulin heavy chain binding protein as glucose-regulated
A:Reference number: A37048; MUID:90130666
A:Accession: A37048
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-655 <K21>
A:Accession: B37048
A:Molecule type: protein
A:Residues: 20-41, 'X', 43-46 <K22>
R:Parfett, C.L.J.; Hofbauer, R.; Brudzynski, K.; Edwards, D.R.; Denhardt, D.T.
Gene 82, 291-303, 1989
A>Title: Differential screening of a cDNA library with cDNA probes amplified in a hetero
A:Reference number: J00094; MUID:90060818
A:Accession: J00094
A:Molecule type: mRNA
A:Residues: 488-590, 'G', 592-595, 'E', 597-655 <PAR>
A:Cross-references: GB:M30779; NID:g193644; PIDN:AAA37742.1; PID:g193645
R:Tillman, J.B.; Moto, P.L.; Walford, R.L.; Splindler, S.R.
Gene 158, 225-229, 1995
A>Title: Structure and regulation of the mouse GRP78 (BiP) promoter by glucose and calci
A:Reference number: I49137; MUID:95331621
A:Accession: I49137
A>Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28 <RES>
A:Cross-references: EMBL:U16277; NID:g829364; PIDN:AAA6734.1; PID:g829365
R:Haas, I.G.; Moos, T.
Proc Natl. Acad. Sci. U.S.A. 85, 2250-2254, 1988
A>Title: cDNA cloning of the immunoglobulin heavy chain binding protein.
A:Reference number: A31934; MUID:88176922
A:Accession: A31934
A:Molecule type: mRNA
A:Residues: 514-589, 'D', 591-595, 'E', 597-655 <HAA>
A:Cross-references: GB:M19351; NID:g192231; PIDN:AAA37315.1; PID:g387113
C:Genetics:
A:Gene: grp78
C:Function:
A:Description: Involved in protein folding and assembling/dissassembling of protein comp
C:Superfamily: heat shock protein 70
C:Keywords: ATP; endoplasmic reticulum; glycoprotein; heterotetramer; molecular chaperon
F.1-19/Domain: signal sequence #status predicted <SIG>

F:20-655/Product: dnaK-type molecular chaperone grp78 #status experimental <MAT>
F:652-655/Region: endoplasmic reticulum retention signal

Query Match 99.3%; Score 3201; DB 2; Length 655;
Best Local Similarity 99.4%; Pred. No. 16-154;
Matches 627; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EDDKKEDVGTAVGIDLTSTYSCGVFRKNGVELIANDGNRTTPSVAFTEBGERLLGDA 61
Db 21 EDDKKEDVGTAVGIDLTSTYSCGVFRKNGVELIANDGNRTTPSVAFTEBGERLLGDA 80
Qy 62 AKNQITSNPENTVPDARLLIGRTMNDPSVQODIKFLPFVYVKKPKPYIYVDIGGGQTKT 121
Db 81 AKNQITSNPENTVPDARLLIGRTMNDPSVQODIKFLPFVYVKKPKPYIYVDIGGGQTKT 140
Qy 122 FAPETISAMVLTAKKKEITAEALYKKTHTAAVTPATFPAFNDQAQAKRQAGTLAGLVAKII 181
Db 141 FAPETISAMVLTAKKKEITAEALYKKTHTAAVTPATFPAFNDQAQAKRQAGTLAGLVAKII 200
Qy 182 NEPTAAALVAGIDRREGEKNIVFDLGGTFDVSLLTINDNGVEVAAFGDTHLGGEDEFD 241
Db 201 NEPTAAALVAGIDRREGEKNILVFDLGGTFDVSLLTINDNGVEVAAFGDTHLGGEDEFD 260
Qy 242 QVMEHFITKLKTKTKGDKVRKDNAAVOKLBRVEKAKRALSSQHQARIEISFYEGEDFS 301
Db 261 QVMEHFITKLKTKTKGDKVRKDNAAVOKLBRVEKAKRALSSQHQARIEISFYEGEDFS 320
Qy 302 ETLTRAKFEELNMDLFRSTMKPVQKVLIEDSDLKKSIDIELVVGSTRIPKIQOLVKEFF 361
Db 321 ETLTRAKFEELNMDLFRSTMKPVQKVLIEDSDLKKSIDIELVVGSTRIPKIQOLVKEFF 380
Qy 362 NCKEPRSGINPDEAVANGAAGVAVSGSDODTDGLVLDVCPVLTLGIEVGVMTKLIPR 421
Db 381 NCKEPRSGINPDEAVANGAAGVAVSGSDODTDGLVLDVCPVLTLGIEVGVMTKLIPR 440
Qy 422 NTVPVPTKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFTDLTGIPAPRGVQIEV 481
Db 441 NTVPVPTKKSQIFSTASDNQPTVTIKYVEGERPLTKDNHLLGTFTDLTGIPAPRGVQIEV 500
Qy 482 TFEIDVNGILVLTAEADGKTGNKNTITNDQNLTPPEELERAVNDAEKFAEDDKLKXRI 541
Db 501 TFEIDVNGILVLTAEADGKTGNKNTITNDQNLTPPEELERAVNDAEKFAEDDKLKXRI 560
Qy 542 DTRNLESEYAVSLANKQIGDKELGKLSSEDKETMEKAVEKEIEMLESHQADIEDFKAK 601
Db 561 DTRNLESEYAVSLANKQIGDKELGKLSSEDKETMEKAVEKEIEMLESHQADIEDFKAK 620
Qy 602 KKELEIYQPIISKLKSGACPPPGEDDTAE 632
Db 621 KKELEIYQPIISKLKSGACPPPGEDDTAE 651

RESULT 3

HRH7GB
dnaK-type molecular chaperone precursor - rat
N:Alternate names: BiP; glucose-regulated 78k protein; Ig heavy chain-binding protein
N:Contains: steroidogenesis-activator polypeptide
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-Jun-1999
C:Accession: A23948; A60134; A26257
R:Munro, S.; Pelham, H.R.B.
Cell 46, 291-300, 1986
A>Title: An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated p
A:Reference number: A23948; MUID:86245075
A:Accession: A23948
A:Molecule type: mRNA
A:Residues: 1-654 <MUN>
A:Cross-references: GB:M14050; NID:g203150; PIDN:AAA40817.1; PID:g203151
R:Pedersen, R.C.; Brownie, A.C.
Science 236, 188-190, 1987
A>Title: Steroidogenesis-activator polypeptide isolated from a rat Leydig cell tumor
A:Reference number: A60134; MUID:87177981

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 14:20:52 ; Search time 23.81 Seconds
(without alignments)
603,931 Million cell updates/sec

Title: US-09-806-955-1

Sequence: 1 MEDKKEDVGVTVGIDLGTT.....AGPPPTGEEDTAEIAHHHHH 639

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

al number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/laa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/laa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/laa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/laa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/laa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/laa/beck11est1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3209	98.0	654	1 US-08-441-139-11	Sequence 11, Appl
2	3168	96.8	666	1 US-08-441-139-16	Sequence 16, Appl
3	2174	66.4	682	1 US-08-441-139-2	Sequence 2, Appl
4	2156	65.9	663	1 US-08-441-139-7	Sequence 7, Appl
5	2077	63.5	646	1 US-08-441-139-14	Sequence 14, Appl
6	2063.5	63.1	679	1 US-08-441-139-5	Sequence 5, Appl
7	1938	59.2	643	4 US-08-97-358B-3	Sequence 3, Appl
8	1547.5	47.3	641	1 US-08-214-583-2	Sequence 2, Appl
9	1474	45.0	649	4 US-08-441-139-4	Sequence 4, Appl
10	1461.5	44.7	607	2 US-09-066-047-5	Sequence 5, Appl
11	1444	44.1	600	6 US-08-472-534-5	Sequence 5, Appl
12	1303.5	39.8	600	6 5240706-1	Sequence 52, Appl
13	1140.5	34.8	339	2 US-08-928-692-52	Sequence 52, Appl
14	941	28.8	187	6 5196523-13	Patent No. 5196523
15	824	25.2	168	1 US-08-441-139-10	Sequence 10, Appl
16	818.5	25.0	315	1 US-08-257-073-7	Sequence 7, Appl
17	750.5	22.9	471	1 US-08-203-905B-2	Sequence 2, Appl
18	726.5	22.2	472	1 US-08-203-905B-14	Sequence 14, Appl
19	642.5	19.6	129	6 5196523-10	Patent No. 5196523
20	607.5	18.6	999	2 US-08-770-301A-3	Sequence 3, Appl
21	607.5	18.6	999	2 US-09-175-581-3	Sequence 3, Appl
22	598	18.3	999	2 US-08-770-301A-1	Sequence 1, Appl
23	598	18.3	999	2 US-08-175-581-1	Sequence 1, Appl
24	528.5	16.1	560	2 US-08-928-692-53	Sequence 53, Appl
25	437.5	13.4	374	2 US-08-928-692-51	Sequence 51, Appl
26	402	12.3	77	6 5196523-7	Patent No. 5196523
27	395	12.1	79	6 5196523-11	Patent No. 5196523

28	367.5	11.2	599	4 US-09-080-983-9	Sequence 9, Appl
29	358	10.9	549	2 US-08-770-544-6	Sequence 6, Appl
30	345.5	10.6	80	1 US-08-464-164-4	Sequence 4, Appl
31	345.5	10.6	80	1 US-08-338-057-4	Sequence 4, Appl
32	345.5	10.6	80	2 US-08-668-416-4	Sequence 4, Appl
33	323	9.9	554	5 PCT-US94-06430-7	Sequence 7, Appl
34	320	9.8	554	4 US-08-591-468-7	Sequence 7, Appl
35	267.5	8.2	714	2 US-08-472-534-3	Sequence 3, Appl
36	264	8.1	55	6 5196523-12	Patent No. 5196523
37	230	7.0	46	6 5196523-8	Patent No. 5196523
38	211	6.4	42	6 5196523-15	Patent No. 5196523
39	192	5.9	38	6 5196523-9	Patent No. 5196523
40	168.5	5.1	1786	4 US-08-973-462-8	Sequence 8, Appl
41	151.5	4.6	1805	1 US-07-853-913-2	Sequence 2, Appl
42	147	4.4	693	4 US-08-235-836C-68	Sequence 68, Appl
43	142.5	4.4	2285	4 US-09-308-375-2	Sequence 2, Appl
44	138	4.2	700	1 US-07-720-589-2	Sequence 2, Appl
45	138	4.2	700	2 US-08-785-190-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-441-139-11
Sequence 11, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltrop, Dr. Karl D.
ATTORNEY: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-11
Query Match 98.0%; Score 3209; DB 1; Length 654;
Best local similarity 99.7%; Pred. No. 6.1e-250;
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 EEDKEDVGVTVGIDGTTYSYGVGFKNGRVEITINDGKNRTTSPYAFPEGERLIGDA 61
DB 20 EEDKEDVGVTVGIDGTTYSYGVGFKNGRVEITINDGKNRTTSPYAFPEGERLIGDA 79
QY 62 AKNOLTSNPENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOWDIGGQTKT 121
DB 80 AKNOLTSNPENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOWDIGGQTKT 139
QY 122 FAPETISAMVLTMKKEFAEYLGKKTTHAVVTPAYFNDARQATKAGTLAGNVNRIT 181
DB 140 FAPETISAMVLTMKKEFAEYLGKKTTHAVVTPAYFNDARQATKAGTLAGNVNRIT 199
QY 182 NEPTAAAIYAGLDKREGEKNILVFDLGSGTSPVSLTTIDNGVFEEVATNGDTHLGSDPD 241
DB 200 NEPTAAAIYAGLDKREGEKNILVFDLGSGTSPVSLTTIDNGVFEEVATNGDTHLGSDPD 259
QY 242 ORVMEHFKIKYKKTKGDKVAKNDRAVQKLRREVEKAKRALSQHQARIEISFEGEDFS 301
DB 260 ORVMEHFKIKYKKTKGDKVAKNDRAVQKLRREVEKAKRALSQHQARIEISFEGEDFS 319
QY 302 ETLTRAFEEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIYLVGSGSTRIPKIQOLVKEFF 361
DB 320 ETLTRAFEEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIYLVGSGSTRIPKIQOLVKEFF 379
QY 362 NGKPSRGINPDEAVAYGAAGVAGVLSGDPDGDVLLDVCPLTGLTFTVGVWTKLIPR 421
DB 380 NGKPSRGINPDEAVAYGAAGVAGVLSGDPDGDVLLDVCPLTGLTFTVGVWTKLIPR 439
QY 422 NTVPVTRKSOIFSTASDNPPTVITIKYEGEERPLTKDNHLLGTFDLTGIPAPRGVPOIEV 481
DB 440 NTVPVTRKSOIFSTASDNPPTVITIKYEGEERPLTKDNHLLGTFDLTGIPAPRGVPOIEV 499
QY 482 TFEIDVNGIILRVTAEDKGKGNKKKITTNDQNRLLPPELEHNVNDAEFAEDKKLKERT 541
DB 500 TFEIDVNGIILRVTAEDKGKGNKKKITTNDQNRLLPPELEHNVNDAEFAEDKKLKERT 559
QY 542 DTRNELESYAVSLKNOIGDEKLGKLSSEDEKTEKAVEEKIEWLESHQADIEDFKAK 601
DB 560 DTRNELESYAVSLKNOIGDEKLGKLSSEDEKTEKAVEEKIEWLESHQADIEDFKAK 619
QY 602 KKELEEVOPILSKLGSAGPPPTGGEEDFAE 632
DB 620 KKELEEVOPILSKLGSAGPPPTGGEEDTSE 650
```

RESULT 2

US-08-441-139-16

Sequence 16, Application US/08441139

Patent No. 5773245

GENERAL INFORMATION:

APPLICANT: Wiltup, Dr. Karl D.

APPLICANT: Robinson, Anne S.

TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/441,139

FILING DATE: 15-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

```
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digilife, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-16
```

```
Query Match 96.8%; Score 3168; DB 1; Length 666;
Best Local Similarity 97.8%; Pred. No. 1,3e-246;
Matches 617; Conservative 11; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 EEDKEDVGVTVGIDGTTYSYGVGFKNGRVEITINDGKNRTTSPYAFPEGERLIGDA 61
DB 32 DEEKEDVGVTVGIDGTTYSYGVGFKNGRVEITINDGKNRTTSPYAFPEGERLIGDA 91
QY 62 AKNOLTSNPENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOWDIGGQTKT 121
DB 92 AKNOLTSNPENTVFDAKRLIGRTWMDPSVQODIKFLPKRYVEKTKPYIOWDIGGQTKT 151
QY 122 FAPETISAMVLTMKKEFAEYLGKKTTHAVVTPAYFNDARQATKAGTLAGNVNRIT 181
DB 152 FAPETISAMVLTMKKEFAEYLGKKTTHAVVTPAYFNDARQATKAGTLAGNVNRIT 211
QY 182 NEPTAAAIYAGLDKREGEKNILVFDLGSGTSPVSLTTIDNGVFEEVATNGDTHLGSDPD 241
DB 212 NEPTAAAIYAGLDKREGEKNILVFDLGSGTSPVSLTTIDNGVFEEVATNGDTHLGSDPD 271
QY 242 ORVMEHFKIKYKKTKGDKVAKNDRAVQKLRREVEKAKRALSQHQARIEISFEGEDFS 301
DB 272 ORVMEHFKIKYKKTKGDKVAKNDRAVQKLRREVEKAKRALSQHQARIEISFEGEDFS 331
QY 302 ETLTRAFEEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIYLVGSGSTRIPKIQOLVKEFF 361
DB 332 ETLTRAFEEELNMDLFRSTMKPVQKYLEDSDLKSDIDEIYLVGSGSTRIPKIQOLVKEFF 391
QY 362 NGKPSRGINPDEAVAYGAAGVAGVLSGDPDGDVLLDVCPLTGLTFTVGVWTKLIPR 421
DB 392 NGKPSRGINPDEAVAYGAAGVAGVLSGDPDGDVLLDVCPLTGLTFTVGVWTKLIPR 451
QY 422 NTVPVTRKSOIFSTASDNPPTVITIKYEGEERPLTKDNHLLGTFDLTGIPAPRGVPOIEV 481
DB 452 NTVPVTRKSOIFSTASDNPPTVITIKYEGEERPLTKDNHLLGTFDLTGIPAPRGVPOIEV 511
QY 482 TFEIDVNGIILRVTAEDKGKGNKKKITTNDQNRLLPPELEHNVNDAEFAEDKKLKERT 541
DB 512 TFEIDVNGIILRVTAEDKGKGNKKKITTNDQNRLLPPELEHNVNDAEFAEDKKLKERT 571
QY 542 DTRNELESYAVSLKNOIGDEKLGKLSSEDEKTEKAVEEKIEWLESHQADIEDFKAK 601
DB 572 DTRNELESYAVSLKNOIGDEKLGKLSSEDEKTEKAVEEKIEWLESHQADIEDFKAK 631
QY 602 KKELEEVOPILSKLGSAGPPPTGGEEDFAE 632
DB 632 KKELEEVOPILSKLGSAGPPPTGGEEDFAE 662
```

RESULT 3

US-08-441-139-2

Sequence 2, Application US/08441139

Patent No. 5773245

<u>DB Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
USPT	recombinant adj immunoglobulin adj heavy adj chain adj binding adj protein	0	<u>L12</u>
USPT	recombinant adj (Ig) adj heavy adj chain adj binding adj protein	0	<u>L11</u>
USPT,PGPB	l3 and (pharmaceutic\$ adj composition\$)	15	<u>L10</u>
USPT,PGPB	l6 and pharmaceutic\$	39	<u>L9</u>
USPT,PGPB	l3 and inflamm\$	14	<u>L8</u>
USPT,PGPB	l3 and pharmaceutic\$	46	<u>L7</u>
USPT,PGPB	l3 and composition\$	72	<u>L6</u>
USPT,PGPB	l4 same composition\$	0	<u>L5</u>
USPT,PGPB	l2 same recombinant\$	4	<u>L4</u>
USPT,PGPB	l2 and recombinant\$	94	<u>L3</u>
USPT,PGPB	((Ig) or immunoglob\$) adj heavy adj chain adj binding adj protein\$	113	<u>L2</u>
USPT,PGPB	recombinant adj heavy adj chain adj binding adj protein\$	0	<u>L1</u>

THIS PAGE BLANK (USPTO)

(FILE 'HOME' ENTERED AT 10:33:09 ON 07 DEC 2001)

FILE 'BIOSIS, MEDLINE, EMBASE, CAPLUS' ENTERED AT 10:33:22 ON 07 DEC 2001
L1 784 S IMMUNOGLOBULIN? (W) HEAVY (W) CHAIN (W) BINDING (W) PROTEIN?
L2 20 S L1 (P) RECOMBINANT
L3 8 DUP REM L2 (12 DUPLICATES REMOVED)

=>

09/806,955

THIS PAGE BLANK (USPTO)

L2 ANSWER 1 OF 20 BIOSIS COPYRIGHT 2001 BIOSIS
TI Effects of co-expressing chaperone BiP on functional antibody production
in the baculovirus system.
AU Hsu, Tsu-An; Eiden, Joseph J.; Bourgarel, Pierre; Meo, Tommaso;
Betenbaugh, Michael J. (1)
SO Protein Expression and Purification, (1994) Vol. 5, No. 6, pp. 595-603.
ISSN: 1046-5928.
AB The assembly pathway of the insect cell *Spodoptera frugiperda* (Sf-9) was
engineered to include expression of the murine chaperone
immunoglobulin heavy chain binding
protein (BiP) using the baculovirus vector. The impact of BiP
coexpression on the production and secretion of functional and soluble
recombinant immunoglobulin IgG levels was evaluated.
Recombinant BiP was found to associate specifically with
immunoglobulins in immunoprecipitation studies. Coinfection of insect
cells with a BiP-containing baculovirus and baculoviruses coding for two
different murine IgG proteins increased intracellular functional antibody
activity levels substantially above the levels observed in the absence of
BiP. Soluble intracellular immunoglobulin levels were found to increase as
well. However, secreted functional antibody levels did not increase
significantly. Also, degradation of heavy chain immunoglobulin in insect
cells was indicated by the accumulation of lower molecular weight
immunoglobulins at 4 days postinfection. Coexpression of light chains
reduced the level of these lower molecular weight immunoglobulins while
BiP coexpression led to enhanced levels. These findings suggest that
coexpressed BiP can increase intracellular soluble and functional antibody
yields but that secretion in the baculovirus-insect cell system must be
limited at some post-translational step.

THIS PAGE BLANK (USPTO)

L2 ANSWER 18 OF 20 CAPLUS COPYRIGHT 2001 ACS
TI Modification of hepatic immunoglobulin heavy chain binding protein
(BiP/Grp78) following exposure to structurally diverse peroxisome
proliferators
AU Witzmann, F. A.; Jarnot, B. M.; Parker, D. N.; Clack, J. W.
SO Fundam. Appl. Toxicol. (1994), 23(1), 1-8
CODEN: FAATDF; ISSN: 0272-0590

AB This investigation was conducted to det. the comparative effect of
structurally diverse peroxisome proliferators (PP) on the two-dimensional
protein pattern of rat liver whole homogenates. Perfluoro-n-decanoic acid
(PFDA), perfluoro-n-octanoic acid (PFOA), clofibrate, and di(2-ethylhexyl)
phthalate (DEHP) are all known to cause the proliferation of hepatic
peroxisomes and the induction of peroxisomal .beta.-oxidative and
microsomal .omega.-oxidative enzymes. To clarify the mechanistic
differences between these compds. with regard to the liver, the authors
examd. the unique patterns of protein alteration produced by in vivo
exposure to them. Following exposure to various doses, whole liver
homogenates were prepd. and sepd. by two-dimensional gel electrophoresis
(2DE) using the ISO-DALT system. Stained gels were digitized and
protein patterns analyzed using the Kepler 2D gel anal. system.

Ig heavy chain binding

protein (BiP), also known as 78-kDa glucose-regulated protein
(Grp78), was identified immunol. and by comigration of **recombinant**
Grp78. BiP is a luminal endoplasmic reticular protein that functions is
the assembly and folding of nascent proteins as they enter the ER. The
present results suggest a selective posttranslational modification of BiP
following PFDA exposure. Single-dose exposure to PFDA was assocd. with a
notable charge modification of BiP that persists up to 30 days. PFOA,
clofibrate, and DEHP had less effect in this regard. The identity of
BiP/Grp78 as the halothane hepatitis-assocd. trifluoroacetylated protein
was also demonstrated. The nature of this PFDA-assocd. protein
modification (reactive metabolite conjugation, abnormal ribosylation, or
phosphorylation) is currently under investigation. The results document
PFDA's unique toxicity as a PP and support the utility of 2D gel anal. in
toxicity testing.

THIS PAGE BLANK (USPTO)

- TI cDNA cloning of the immunoglobulin heavy chain binding protein
AU Haas, Ingrid G.; Meo, Tommaso
SO Proc. Natl. Acad. Sci. U. S. A. (1988), 85(7), 2250-4
CODEN: PNASA6; ISSN: 0027-8424
- AB A cDNA library was constructed from size-fractionated poly(A)+ RNA prepd. from a murine pre-B-cell hybridoma expressing high levels of **Ig heavy-chain-binding protein (BiP)** and μ heavy chains. Transformed bacterial colonies were screened for **recombinant** plasmids contg. cDNA coding for BiP by hybrid-selected mRNA translation. A clone, pMBiP, contg. a 736-base-pair insert was shown to encode the protein. Translation in vitro of hybridoma mRNA selected by hybridization to the pMBiP cDNA yielded a single polypeptide of BiP-like size. The authenticity of this mRNA was verified by comparing the peptides obtained by the limited proteolysis of its in vitro translation product with those obtained from the in vivo produced BiP. Likewise, the authenticity of the cDNA insert was verified by an RNase A protection assay of heteroduplex mols. obtained by annealing a uniformly labeled single-strand copy of the cDNA clone with the same mRNA selected by hybridization and tested by translation. The nucleotide sequence of this clone was used to deduce the C-terminal 142 amino acids of BiP and to establish its kinship with the 70-kDa heat-shock protein family. The finding of a single copy of the BiP gene in DNA blots of mouse and rat implies that the BiP-related RNA transcripts constitutively expressed in various murine tissues and cell lines are indeed products of the same gene. These findings imply that BiP plays a more general role than previously anticipated on the basis of the discovery of its assocn. with Ig heavy chains.

THIS PAGE BLANK (USPTO)

L2 ANSWER 8 OF 20 MEDLINE

TI Identification of immunoglobulin heavy chain binding protein as glucose-regulated protein 78 on the basis of amino acid sequence, immunological cross-reactivity, and functional activity.

AU Kozutsumi Y; Normington K; Press E; Slaughter C; Sambrook J; Gething M J

SO JOURNAL OF CELL SCIENCE. SUPPLEMENT, (1989) 11 115-37.

Journal code: HNG; 8502898. ISSN: 0269-3518.

AB **Immunoglobulin heavy chain binding**

protein (BiP) associates transiently with various proteins destined for the secretory pathway. To investigate the relationship between BiP and the 78K (K = 10(3) Mr) glucose-regulated protein (GRP78), we have determined a partial amino acid sequence of purified mouse BiP and isolated and sequenced a full-length cDNA clone encoding mouse GRP78. The 26 amino-terminal residues of the mature BiP protein are identical to a sequence of amino acids located near the start of the open reading frame encoding GRP78. A polyclonal antiserum raised against mouse GRP78 protein expressed in bacteria from the cloned GRP78 cDNA could immunoprecipitate complexes consisting of BiP and unfolded forms of immunoglobulin heavy chains. Furthermore, a monoclonal antibody raised against mouse BiP immunoprecipitated mouse GRP78 expressed in monkey CV-1 cells from an SV40-GRP78 **recombinant** vector. Finally, like the endogenous BiP of simian cells, mouse GRP78 associated with malfolded, non-glycosylated forms of influenza hemagglutinin (HA) when GRP78 and HA were co-expressed from SV40 vectors in CV-1 cells. These studies confirm that BiP is identical to GRP78. Comparison of the nucleic acid and deduced amino acid sequence of mouse GRP78 with those of other rodent and human GRP78s revealed an extremely high degree of sequence identity. BiP/GRP78 is closely related (approximately 60% identity) to the cytoplasmic 70K heat-shock proteins. Surprisingly, the carboxy-terminal 29 amino acids of BiP/GRP78, which are not conserved in HSP70 proteins, are almost identical in sequence to the steroidogenesis activator peptide found in the cytoplasm of rat Leydig tumor cells. Possible relationships between these polypeptides are discussed.

THIS PAGE BLANK (USPTO)